

IOWA STATE UNIVERSITY

Digital Repository

Plant Pathology and Microbiology Publications

Plant Pathology and Microbiology

2018

Rpp1 encodes a ULP1-NBS-LRR protein that controls immunity to *Phakopsora pachyrhizi* in soybean

Kerry F. Pedley

United States Department of Agriculture

Ajay K. Pandey

United States Department of Agriculture

Amy Ruck

United States Department of Agriculture

Lori M. Lincoln

United States Department of Agriculture

Steven A. Whitham

Follow this and additional works at: https://lib.dr.iastate.edu/plantpath_pubs

Iowa State University, swhitham@iastate.edu

 Part of the [Agricultural Science Commons](#), [Agriculture Commons](#), [Plant Breeding and Genetics Commons](#), and the [Plant Pathology Commons](#)

See next page for additional authors

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/plantpath_pubs/260. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Plant Pathology and Microbiology at Iowa State University Digital Repository. It has been accepted for inclusion in Plant Pathology and Microbiology Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Rpp1 encodes a ULP1-NBS-LRR protein that controls immunity to *Phakopsora pachyrhizi* in soybean

Abstract

Phakopsora pachyrhizi is the causal agent of Asian soybean rust. Susceptible soybean plants infected by virulent isolates of *P. pachyrhizi* are characterized by tan-colored lesions and erumpent uredinia on the leaf surface. Germplasm screening and genetic analyses have led to the identification of seven loci, Rpp1 – Rpp7, that provide varying degrees of resistance to *P. pachyrhizi* (Rpp). Two genes, Rpp1 and Rpp1b, map to the same region on soybean chromosome 18. Rpp1 is unique among the Rpp genes in that it confers an immune response (IR) to avirulent *P. pachyrhizi* isolates. The IR is characterized by a lack of visible symptoms, whereas resistance provided by Rpp1b – Rpp7 results in red-brown foliar lesions. Rpp1 maps to a region spanning approximately 150 Kb on chromosome 18 between markers Sct_187 and Sat_064 in L85-2378 (Rpp1), an isoline developed from Williams 82 and PI 200492 (Rpp1). To identify Rpp1, we constructed a bacterial artificial chromosome (BAC) library from soybean accession PI 200492. Sequencing of the Rpp1 locus identified three homologous nucleotide binding site-leucine rich repeat (NBS-LRR) candidate resistance genes between Sct_187 and Sat_064. Each candidate gene is also predicted to encode an N-terminal ubiquitin-like protease 1 (ULP1) domain. Co-silencing of the Rpp1 candidates abrogated the immune response in the Rpp1 resistant soybean accession PI 200492, indicating that Rpp1 is a ULP1-NBS-LRR protein and plays a key role in the IR.

Keywords

Plant immunity, soybean rust, sumoylation, integrated decoy

Disciplines

Agricultural Science | Agriculture | Plant Breeding and Genetics | Plant Pathology

Comments

This is a manuscript of an article published as Pedley, Kerry F., Ajay K. Pandey, Amy Ruck, Lori M. Lincoln, Steven A. Whitham, and Michelle A. Graham. "Rpp1 encodes a ULP1-NBS-LRR protein that controls immunity to *Phakopsora pachyrhizi* in soybean." *Molecular Plant-Microbe Interactions* (2018). doi: [10.1094/MPMI-07-18-0198-FI](https://doi.org/10.1094/MPMI-07-18-0198-FI).

Rights

Works produced by employees of the U.S. Government as part of their official duties are not copyrighted within the U.S. The content of this document is not copyrighted.

Authors

Kerry F. Pedley, Ajay K. Pandey, Amy Ruck, Lori M. Lincoln, Steven A. Whitham, and Michelle A. Graham

1 ***Rpp1* encodes a ULP1-NBS-LRR protein that controls immunity to *Phakopsora***
2 ***pachyrhizi* in soybean**

3

4 Kerry F. Pedley,^{1,†} Ajay K. Pandey,^{1,3} Amy Ruck,¹ Lori M. Lincoln,² Steven A. Whitham,³
5 and Michelle A. Graham^{2,†}

6

7 ¹United States Department of Agriculture–Agricultural Research Service (USDA-ARS),
8 Foreign Disease-Weed Science Research Unit, Ft. Detrick, MD 21702; ²USDA-ARS, Corn
9 Insects and Crop Genetics Research Unit, Ames, IA 50011; ³Iowa State University,
10 Department of Plant Pathology and Microbiology, Ames, IA 50011

11

12 Current address of A. K. Pandey: National Agri-Food Biotechnology Institute, Sector 81,
13 Mohali, India 140306

14

15 †Corresponding authors: K. F. Pedley; E-mail: Kerry.Pedley@ars.usda.gov and
16 M. A. Graham; E-mail: Michelle.Graham@ars.usda.gov

17

18 Keywords: Plant immunity, soybean rust, sumoylation, integrated decoy

19

20

21 **Commercial Endorsement Disclaimer** – Mention of trade names or commercial
 22 products in this publication is solely for the purpose of providing specific information
 23 and does not imply recommendation or endorsement by the U.S. Department of
 24 Agriculture.

25
 26 **Equal Opportunity/Non-Discrimination Statement** – The U.S. Department of
 27 Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis
 28 of race, color, national origin, age, disability, and where applicable, sex, marital status,
 29 familial status, parental status, religion, sexual orientation, genetic information, political
 30 beliefs, reprisal, or because all or a part of an individual's income is derived from any
 31 public assistance program. (Not all prohibited bases apply to all programs.) Persons
 32 with disabilities who require alternative means for communication of program
 33 information (Braille, large print, audiotope, etc.) should contact USDA's TARGET Center
 34 at (202) 720-2600 (voice and TDD). To file a complaint of discrimination write to USDA,
 35 Director, Office of Civil Rights, 1400 Independence Avenue, SW, Washington, DC 20250-
 36 9410 or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal
 37 opportunity provider and employer.

38
 39

ABSTRACT

Phakopsora pachyrhizi is the causal agent of Asian soybean rust. Susceptible soybean plants infected by virulent isolates of *P. pachyrhizi* are characterized by tan-colored lesions and erumpent uredinia on the leaf surface. Germplasm screening and genetic analyses have led to the identification of seven loci, *Rpp1* – *Rpp7*, that provide varying degrees of resistance to *P. pachyrhizi* (*Rpp*). Two genes, *Rpp1* and *Rpp1b*, map to the same region on soybean chromosome 18. *Rpp1* is unique among the *Rpp* genes in that it confers an immune response (IR) to avirulent *P. pachyrhizi* isolates. The IR is characterized by a lack of visible symptoms, whereas resistance provided by *Rpp1b* – *Rpp7* results in red-brown foliar lesions. *Rpp1* maps to a region spanning approximately 150 Kb on chromosome 18 between markers Sct_187 and Sat_064 in L85-2378 (*Rpp1*), an isolate developed from Williams 82 and PI 200492 (*Rpp1*). To identify *Rpp1*, we constructed a bacterial artificial chromosome (BAC) library from soybean accession PI 200492. Sequencing of the *Rpp1* locus identified three homologous nucleotide binding site-leucine rich repeat (NBS-LRR) candidate resistance genes between Sct_187 and Sat_064. Each candidate gene is also predicted to encode an N-terminal ubiquitin-like protease 1 (ULP1) domain. Co-silencing of the *Rpp1* candidates abrogated the immune response in the *Rpp1* resistant soybean accession PI 200492, indicating that *Rpp1* is a ULP1-NBS-LRR protein and plays a key role in the IR.

60 INTRODUCTION

61 Plants are a potential source of sugars and other nutrients to microbes in the
 62 environment. Thus, it is not surprising that many bacteria and fungi have evolved
 63 mechanisms to overcome the preformed structural and chemical barriers plants use to
 64 protect these resources. As a result, plants have evolved additional systems for
 65 detecting would-be pathogens and mounting defenses to thwart infection. Collectively
 66 these systems and the underlying mechanisms are referred to as plant immunity.

67 It is now well-established that plants utilize two distinct, but overlapping
 68 immune responses (Jones and Dangl, 2006). The recognition of microbe- or pathogen-
 69 associated molecular patterns (MAMPs or PAMPs) activates MAMP- or PAMP-triggered
 70 immunity (MTI or PTI) (Jones and Dangl, 2006; Dodds and Rathjen, 2010). To counter
 71 MTI, pathogens deliver effector proteins into the plant cell that interfere with signaling
 72 and other downstream defense responses, resulting in effector-triggered susceptibility
 73 (ETS). Recognition or detection of microbial effectors by plant resistance proteins (R
 74 proteins) elicits effector-triggered immunity (ETI), which often culminates with a rapid,
 75 localized form of cell death known as the hypersensitive response (HR) (Jones and Dangl,
 76 2006; Boller and Felix, 2009; Dodds and Rathjen, 2010). The HR functions to limit the
 77 ability of the pathogen to spread beyond the point of attack. Due to the dramatic and
 78 irreversible nature of the HR, it is essential that plants tightly regulate ETI in the absence
 79 of a bona fide infection.

80 The majority of plant R proteins contain nucleotide-binding site leucine-rich
 81 repeat (NBS-LRR) domains and are encoded by plant resistance genes (*R* genes) (Martin

et al., 2003). Many NBS-LRR proteins can be further divided into three functional classes based upon the presence of either Toll/interleukin-1 receptor (TIR), coiled coil (CC), or resistance to powdery mildew8 (RPW8) domains located at the amino terminus (Jones et al., 2016; Shao et al., 2016). NBS-LRR proteins are involved in pathogen detection with the TIR-NBS-LRR and CC-NBS-LLR classes operating though distinct, yet overlapping signaling pathways (Aarts et al., 1998; Elmore et al., 2011). Although many aspects of R protein function remain to be elucidated, there is evidence for both direct and indirect interactions of NBS-LRR proteins with microbial effectors, which initiate signaling cascades that trigger defense responses (Jones and Dangl, 2006; Dodds and Rathjen, 2010; Jones et al., 2016).

Soybean rust (SBR) is an aggressive foliar disease of soybean (*Glycine max*) and other legumes caused by the obligate biotrophic fungus, *Phakopsora pachyrhizi*. Three distinct reaction phenotypes occur on soybean in response to *P. pachyrhizi* infection (Bromfield, 1984), and the outcome of the interaction appears to follow the gene-for-gene resistance model (Flor, 1946). Susceptible soybean plants infected with virulent isolates of *P. pachyrhizi* develop tan-colored lesions (TAN phenotype) with abundant, sporulating uredinia on the leaf surface. Resistant plants are characterized by either a complete lack of visible symptoms, termed the immune response (IR), or by the appearance of reddish-brown lesions (RB reaction) in response to avirulent isolates. The IR effectively prevents reproduction of the pathogen and provides complete resistance (Miles et al., 2003). The RB reaction varies with respect to the pathogen's ability to form

103 uredinia and sporulate, and is considered incomplete resistance (Miles et al., 2003;
104 Bonde et al., 2006).

105 While the identity of the effectors or other determinants that contribute to the
106 recognition of soybean rust isolates remain unknown, germplasm screening efforts have
107 identified several sources of resistance to *P. pachyrhizi* (*Rpp*) that map to seven loci
108 (*Rpp1* – *Rpp7*; Hyten et al., 2007; Garcia et al., 2008; Silva et al., 2008; Hyten et al., 2009;
109 Li et al., 2012; Childs et al., 2017). Two genes, *Rpp1* (Hyten et al., 2007) and *Rpp1b*
110 (Chakraborty et al., 2009), map to an overlapping region on chromosome 18. *Rpp4* (Silva
111 et al., 2008) and *Rpp6* (Li et al., 2012) also map to chromosome 18, but to loci distinct
112 from *Rpp1*. The remaining genes, *Rpp2* (Silva et al., 2008), *Rpp3* (Hyten et al., 2009), the
113 dominant and recessive alleles of *Rpp5* (Garcia et al., 2008), and *Rpp7* (Childs et al.,
114 2017) map to chromosomes 16, 6, 3, and 19, respectively. With the exception of *Rpp1*,
115 all of the known *Rpp* genes condition a RB reaction. *Rpp1*, but not *Rpp1b*, conditions an
116 IR in response to avirulent isolates of *P. pachyrhizi*.

117 The soybean *Rpp* genes have yet to be cloned, but a candidate for *Rpp4* was
118 identified in the resistant accession PI 459025B (*Rpp4*) using a combination of DNA
119 sequencing, gene expression analysis, and virus-induced gene silencing (VIGS) (Meyer et
120 al., 2009). Sequencing the *Rpp4* locus in the susceptible cultivar Williams 82 revealed
121 the presence of three CC-NBS-LRR *R* genes with similarity to the RGC2 *R* gene family in
122 lettuce. PI 459025B contains alleles of these genes not present in susceptible plants, and
123 one allele (*Rpp4C4*) was predominantly expressed in resistant plants before and after
124 challenge with an avirulent *P. pachyrhizi* isolate. Gene silencing resulted in the loss of

resistance in PI 459025B, indicating that *Rpp4* is a CC-NBS-LRR *R* gene within this gene cluster. To understand downstream genes contributing to *Rpp4*-mediated resistance, Morales et al. (2013) used microarray analyses to compare gene expression in *Rpp4*-silenced and control plants two weeks after infection.

In this study, we employed a similar approach to identify the source of resistance at the *Rpp1* locus involved in the IR. We screened a BAC library developed from the *Rpp1* resistant accession PI 200492 with a variety of markers targeting the *Rpp1* locus mapped by Hyten et al. (2007). Sequencing of the BACs revealed the presence of three novel *R* gene candidates between markers Sct_187 and Sat_064, which define the *Rpp1* locus. Co-silencing of the candidate genes using VIGS confirmed a role in resistance to *P. pachyrhizi*. RNA-seq of VIGS plants identified genes involved in the *Rpp1*-mediated defense network. This work shows that *Rpp1* is an NBS-LRR protein with a novel Ubiquitin-like-specific protease 1 (ULP1) domain.

RESULTS

Identification of candidate *Rpp1* genes

To identify candidate *Rpp1* genes, a BAC library was constructed using DNA isolated from the resistant genotype (PI 200492). Markers defining the *Rpp1* locus (Sct_187 and Sat_064) (Hyten et al., 2007) were used to identify the corresponding region in the Williams 82 genome sequence (Gm18:56,182,230 to 56,333,803). Primers for BAC library screening were developed from genes Glyma.18g280900 and Glyma.18g282700, which have relatively few homologs within the soybean genome

(Supplemental Table S1). BAC library screening with these primers identified BACs PI_200492_2G03 and PI_200492_2E10 (Figure 1). Screening with BAC-end primers developed from PI_200492_2G03 and PI_200492_2E10 identified BACs PI_200492_1C06 and PI_200492_1A02. Paired end sequencing of 691, 402, 935 and 754 subclones from BACs PI_200492_1C06, PI_200492_2G03, PI_200492_1A02 and PI_200492_2E10, respectively, was used to generate a 324,316 bp contig (GenBank Accession MH590229) extending past the *Rpp1* locus in both directions. BACs were assembled individually prior to complete contig assembly. BAC coverage ranged from 6.2X for PI_200492_1C06 to 10.0X for BAC PI_200492_1A02.

BLASTN comparisons against primary transcripts in the Williams 82 reference genome identified 31 orthologous genes in the PI 200492 *Rpp1* contig (Glyma.18G280200-Glyma.18G282800, Glyma.18G283100, Glyma.18G283200, Glyma.U008900, Glyma.U008800). Of these, eight genes had homology to the NBS-LRR family of disease resistance genes. The genes are abbreviated as *R1* – *R8*, corresponding to Glyma.18G280300, Glyma.18G280400, Glyma.18G281500, Glyma.18G281600, Glyma.18G281700, Glyma.18G283200, Glyma.U008800 and Glyma.U008900.

Correspondence was determined by resistance gene order relative to the Williams 82 genome sequence. Based on multiple sequence alignments, the presence of conserved domains and the presence or absence of introns, the eight candidate *R* genes could be divided into two classes (Figure 2). Genes *R1* – *R5*, contained NB-ARC (IPR002182), Winged helix-turn-helix (IPR011991), and LRR (IPR032675) domains distinct from those in *R6*, *R7* and *R8*. *R1* and *R3* – *R5* also contained a novel ULP1 protease domain

(IPR003653). *R2* lacked the ULP1 protease domain and contained several in-frame stop codons, making it likely a pseudogene. *R1* and *R3 – R5*, contained three predicted introns, while *R2* contained a single predicted intron. *R6* and *R7* contained an additional coiled-coil domain and lacked introns entirely. *R8* encoded a partial NBS domain and is also a likely pseudogene. Three *R* genes (*R3*, *R4* and *R5*), were located between markers Sct_187 and Sat_064 (Figure 1), which define the *Rpp1* locus (Hyten et al., 2007).

VIGS of the *Rpp1* candidate genes *R3*, *R4* and *R5* compromises immunity

To determine if *R3*, *R4* or *R5* played a role in immunity, a VIGS construct was developed to silence their expression. If any of these genes were required for the IR, we expected visible lesions indicative of a compromised defense response to develop on silenced *Rpp1*-resistant plants challenged with an avirulent isolate of *P. pachyrhizi*. Since the genes shared 93 to 96% nucleotide identity, this prohibited the silencing of each gene individually. Therefore an identical 234 bp portion of the central NBS shared by all three candidate genes was used to generate a single VIGS construct for silencing. Aside from the viral symptoms associated with BPMV, propagation of the virus containing the *Rpp1* insert (BPMV:*Rpp1*) did not result in any morphological abnormalities in susceptible or resistant plants, which were visually indistinguishable from control BPMV:*GFP*-infected plants. For VIGS experiments, *Rpp1*-silenced and control plants were challenged with *P. pachyrhizi* isolate LA04-1 4 weeks following the introduction of BPMV:*Rpp1* by rub inoculation. The *Rpp1*-silenced plants exhibited visible RB lesions with limited uredinia formation 14 d after inoculation with isolate LA04-1 (Figure 3). By contrast, resistance was not affected in plants containing the BPMV:*GFP* recombinant

191 virus and no visible symptoms were observed. To confirm silencing, mRNA from three
 192 independent biological replicates was measured by qRT-PCR. The *Rpp1* candidate gene
 193 transcript levels were reduced 1.67 ± 0.19 -fold in BPMV:*Rpp1*-infected plants compared
 194 to BPMV:*GFP*-infected plants. To verify that *Rpp1*-mediated immunity was compromised
 195 in the silenced plants, the accumulation of *P. pachyrhizi* α -tubulin transcript was used as
 196 a measurement of fungal growth. qRT-PCR performed with cDNA derived from three
 197 independent biological replicates from BPMV:*GFP*-infected plants did not produce a
 198 positive Ct value through 40 cycles. In contrast, three independent biological replicates
 199 of BPMV:*Rpp1*-infected plants yielded an average Ct value of 36.05 ± 0.56 , indicating the
 200 *Rpp1*-silencing allowed fungal accumulation to occur.

201 **R4 is the most highly expressed candidate gene in PI 200492, regardless of infection**

202 In order to determine which of the candidate *R* genes contribute to *P. pachyrhizi*
 203 resistance, a region of the NBS containing nucleotide differences that could be used to
 204 distinguish each of the genes was PCR amplified from cDNA, cloned and sequenced.
 205 Primers were designed from flanking regions that were absolutely conserved across all
 206 five genes in both PI 200492 and Williams 82 to reduce differences in amplification
 207 efficiency (Supplemental Figure S1). Prior to expression analyses, primers were tested
 208 on genomic DNA of PI 200492 and Williams 82 to determine amplification efficiency for
 209 each gene (Supplemental Table S2). In PI 200492 plants grown under standard
 210 conditions in the absence of the pathogen, 7.5, 0, 6.2, 65.3 and 21 % of clones
 211 corresponded to *R1* – *R5*, respectively, suggesting *R4* is the predominantly expressed
 212 gene in PI 200492 and confirming *R2* as a pseudogene. In contrast, in susceptible

Williams 82 plants 2.6, 2.5, 2.7, 10.5 and 81.8 % of clones corresponded to *R1* – *R5*, respectively, suggesting *R5* is the predominantly expressed gene in Williams 82. In order to test how PI 200492 responded to *P. pachyrhizi* infection, we repeated the experiment collecting samples 24 and 72 hours after inoculation or mock-inoculation. Again, *R4* was the predominantly expressed gene (50.7 to 63.7 % of clones), followed by *R5* (14.9 to 18.4 % of clones). *R1* and *R4*, were slightly induced by *P. pachyrhizi* inoculation at both time points, while *R3* was repressed. *R5* was slightly repressed by *P. pachyrhizi* inoculation at 24 hours, but induced by 72 hours.

221 **The Rpp1 ULP domain is a functional protease**

To determine if the ULP1 domain was functional, we used the ULP1 protein domains from *R1*, *R3*, *R4*, *R5*, Glyma.13G256800 and Glyma.15G058100 for BLASTP comparisons to all proteins in the *Arabidopsis* genome (TAIR version 10). Using an E-value cutoff of 10E-4, we identified seven predicted proteins (ULP1B, ELS1, ASP1, At4G33620, OST1 and OTS2). Sequence alignment of the ULP1 domains encoded by the *Rpp1* candidate genes, the *Arabidopsis* ULP proteins and the two ULPs from *Saccharomyces cerevisiae* confirmed the presence of the catalytic triad residues (His-Asp-Cys, Supplemental Figure S2) required for protease activity. To test for functionality, the ULP1 domains of *R3*, *R4* and *R5* were used to complement *S. cerevisiae* strains that contained mutations in either of the two yeast small ubiquitin-like modifier (SUMO) protease genes, *Ulp1* or *Ulp2*. *Ulp1* has several functions in yeast including the processing of Smt3 (SUMO) precursor peptides and the removal of Smt3 from post-translational conjugates (Li and Hochstrasser, 1999). Strains lacking a functional *Ulp1*

allele are not viable due to their inability to progress through the cell cycle, but *ulp1* temperature sensitive (ts) alleles have been isolated (Li and Hochstrasser, 1999). A *ulp1-ts* allele permits growth at 30°C, but growth at 37°C is restricted. The second *S. cerevisiae* SUMO protease, Ulp2, also functions to remove Smt3 from other proteins, but is functionally distinct from Ulp1 (Li and Hochstrasser, 2000). Null *Ulp2* mutants are viable, but exhibit a pleiotropic phenotype that includes temperature-sensitive growth (Li and Hochstrasser, 2000). We expressed the soybean ULP1 domains using a strong galactose inducible promoter. Additionally, we used a truncated galactose-inducible promoter for weaker expression, in case the expressed soybean gene fragment was toxic in yeast. Control cells transformed with the empty vectors were unable to grow at 37°C on media containing glucose or galactose. The *R5* ULP1 domain was able to restore growth in the *ulp1-ts* when expressed at high levels and the *R4* ULP1 domain complemented the *ulp2* mutant cells at both high and low expression levels (Figure 4). Conversely, expression of the *R4* ULP1 domain encoding a mutation in the catalytic triad, (C310S), failed to complement the *ulp2* mutant (Supplemental Figure S3).

250 **Evolution of the *Rpp1* locus across legumes**

251 To examine the evolution of the *Rpp1* candidate genes across legumes, we took
252 advantage of the Legume Information System Genomic Context Viewer
253 (https://legumeinfo.org/lis_context_viewer/; Dash et al. (2016)). Since *R3*, *R4* and *R5*
254 were located within the mapped *Rpp1* locus and were closely related to *R1* and *R2*, we
255 focused our analyses on the region containing *R1* through *R5* (Gm18 from 56.10 to 56.30
256 MB, version Wm82.a2.v1). We identified the homeologous region on soybean Gm8,

representing a genome duplication event, and syntenic regions from *P. vulgaris* (common bean), *A. duranensis* and *A. ipaensis* (the diploid ancestors of cultivated peanut), *C. cajan* (Pigeon Pea), *L. angustifolius* (Narrow-leaved lupine), *L. japonicus* (Birdsfoot treefoil) and *M. truncatula* (Barrel Medic) (Figure 5). Colinearity was most conserved in legume species most closely related to soybean. We identified two genes, an ethylene responsive transcription factor and a homeodomain transcription factor, that flanked the *Rpp1* candidate genes *R3*, *R4* and *R5* on Gm18 and were conserved across all eight species and the homeologous region on Gm8. Only Gm18 contained *R* genes between these flanking genes. InterProScan analyses of all *R* proteins from across the entire region in all species confirmed that only the *R* genes in the Williams 82 Gm18 reference genome contained the novel ULP1 protease domain. Further, BLAST searches failed to identify any additional *R* proteins with ULP1 domains from the predicted proteins of any of these species. Taken together, this suggests that the presence of *R* genes in the *Rpp1* locus and the addition of the ULP1 protease domain is relatively new, following the separation of soybean and its closest sequenced relative, *P. vulgaris*. Similarly, Meyer et al. (2009) found that *R* genes were absent from the region homeologous to the *Rpp4* Asian soybean rust resistant locus. Further, only one additional *Rpp4* homolog could be found elsewhere in the soybean genome.

In order to understand how the novel ULP1 protease domain was incorporated into the *Rpp1* candidate genes, we used BLASTN to compare the ULP1 domains from *R1*, *R3*, *R4* and *R5* to all predicted transcripts in the Williams 82 reference genome. This search identified best reciprocal matches to two homeologous genes Glyma.13G256800

($E < 0.0$) and Glyma.15G058100 ($E < 10E-17$), both lacking any R protein signatures (Figure 6). While the ULP1 domain in the *Rpp1* candidates corresponded to a single exon, it corresponded to six exons in Glyma.13G256800 and Glyma.15G058100, which have 16 and 10 exons, respectively. Li et al. (2017) identified 13 ULP homologs in the soybean genome, each containing a minimum of four exons. The lack of introns in the ULP1 domain of the *Rpp1* R genes suggests the ULP1 domain was likely inserted into an ancestral R gene at the *Rpp1* locus by a retrotransposition event, likely within the first intron. Genes flanking a retrotransposon can be duplicated and transposed by readthrough transcription from the retroelement (Hoen et al., 2006). We took advantage of the soybean transposable element database, SoyTEdb (Du et al., 2010), to search for transposable elements within all R gene sequences of PI 200492 and Williams 82. We identified a Copia LTR retrotransposon (RLC_Gmr24_Gm10-Gm18, $E=0$) in intron 1 of *R1* in both cultivars (Figure 6). The element immediately preceded the ULP1 domain of *R1*. Since the element is truncated relative to RLC-GMr24 (2035 of 2591 bases) and lacked typical long terminal repeats, it is likely no longer functional.

294 **RNA-seq of *Rpp1* Silenced Plants**

295 Previous experiments silencing *Rpp4* in PI 459025B resulted in the development
296 of tan lesions with fully sporulating uredinia (Meyer et al., 2009). Microarray analyses
297 of *Rpp4*-silenced plants revealed that genes normally induced during *Rpp4*-mediated
298 resistance were repressed by *Rpp4* silencing (Morales et al., 2013). In contrast, *Rpp1*
299 silencing resulted in plants with RB lesions and limited uredinia formation. However,
300 inoculation of PI 200492 with compatible *P. pachyrhizi* isolates results in susceptible

301 TAN lesions (Miles et al., 2011). This suggests that silencing *Rpp1* compromised
302 immunity to *P. pachyrhizi*, but not resistance.

303 To evaluate the effect of *Rpp1* silencing, the RNA used to confirm silencing was
304 also used for RNA-seq analysis. By comparing gene expression between PI 200492
305 BPMV:*Rpp1*-infected plants and BMPV:*GFP*-infected plants, each infected with *P.*
306 *pachyrhizi* isolate LA04–1, we hoped to identify downstream components of the *Rpp1*
307 signaling pathway responsible for resistance to *P. pachyrhizi*. Sequences for each of the
308 samples are available from the National Center for Biotechnology Sequence Read
309 Archive (<https://www.ncbi.nlm.nih.gov/sra>), BioProject Accession PRJNA479513.

310 Using a false discovery rate (FDR) < 0.001, we identified 1,211 genes induced by
311 *Rpp1* silencing and 2,566 genes repressed by *Rpp1* silencing (Fold Change>2,
312 Supplemental Table S3 and Fold change <-2, Supplemental Table S4). To better
313 understand the biological processes affected by *Rpp1* silencing, we used gene ontology
314 (GO) terms to group differentially expressed genes by function. We identified 32 and 72
315 GO terms significantly (Corrected P<0.05) overrepresented among genes induced and
316 repressed by *Rpp1* silencing, respectively (Supplemental Table S5). Among the induced
317 genes, we identified significant GO terms associated with defense (defense response,
318 defense response to fungus, regulation of plant hypersensitive response, respiratory
319 burst involved in defense response, flavonoid biosynthesis, positive regulation of
320 flavonoid biosynthesis and phenylpropanoid biosynthesis), transport (amino acid,
321 nitrate, proline, adenine, guanine and oligopeptide transport, amino acid import and
322 phloem sucrose loading), metabolism (chlorophyll, lipoate, vitamin and sulfur amino

323 acid, oxidoreduction coenzyme and secondary metabolism and regulation of lipid
 324 metabolism) and biosynthesis (coenzyme, oxylipin, fat-soluble vitamin and sulfur
 325 compound biosynthesis). GO terms overrepresented among *Rpp1*-silencing repressed
 326 genes included those associated with photosynthesis (photosynthesis, light reaction,
 327 light harvesting, response to far red and red light, electron transport and others),
 328 growth (growth, cell tip growth, multidimensional growth, regulation of cell size and
 329 meristem growth), defense response (incompatible interaction), abiotic stress responses
 330 (response to cold, dessication, sucrose and temperature stimulus) and biosynthesis
 331 (biosynthesis of chlorophyll, brassinosteroid, lignin, cutin, carbohydrate and lipids). This
 332 suggests that *Rpp1* silencing repressed photosynthetic processes, abiotic stress
 333 responses and growth while inducing defense, similar to resistant responses governed
 334 by *Rpp2* (van de Mortel et al., 2007), *Rpp3* (Schneider et al., 2011) and *Rpp4* (Morales et
 335 al., 2013).

336 Morales et al. (2013), identified 54 probes unique to resistance responses
 337 governed by *Rpp2*, *Rpp3* and *Rpp4*. Using the SoyBase Gene Model Correspondence
 338 Lookup (<https://www.soybase.org/correspondence/>), these correspond to 42 genes in
 339 the current genome assembly. Of these, 11 were also differentially expressed in
 340 response to *Rpp1* silencing. Morales et al. (2013) also used microarray analysis to
 341 identify genes differentially expressed in response to *Rpp4* silencing. Of the 260 genes
 342 differentially expressed in response to *Rpp4* silencing, 97 were also differentially
 343 expressed in response to *Rpp1* silencing. Of these, 89 genes (92%) were repressed by

344 *Rpp4*-silencing. In contrast, 62 genes (70%) were induced by *Rpp1* silencing, again
345 suggesting that *Rpp1* silencing compromised the IR but not defense.

346 To identify the transcription factors responding to *Rpp1* silencing, we took
347 advantage of the SoyDB Transcription Factor database (Wang et al., 2010). We identified
348 348 differentially expressed transcription factors, representing 39 transcription factor
349 families (Supplemental Table S6). Of these, WRKY, NAC and PLATZ transcription factors
350 were significantly over represented among genes differentially expressed in response to
351 *Rpp1* silencing (Corrected-P <0.01). Of the 43 WRKY transcription factors identified, 12
352 were repressed (log2FC from -1.73 to -3.32) by *Rpp1* silencing, including homologs of
353 AtWRKYs 6, 15, 40, 50, 51 and 70. Induced WRKYs (log2FC from 1.39 to 6.94) included
354 homologs of AtWRKYs 6, 23, 28, 30, 33, 40, 41, 42, 57, 65, 72 and 75. WRKYs are
355 associated with abiotic stress tolerance (WRKYs 6, 30, 33, 40, 41, 42, 57, 72, 75) and
356 biotic stress responses (WRKYs 6, 28, 33, 40, 72) (reviewed by (Bakshi and Oelmüller,
357 2014; Phukan et al., 2016)). Glyma.13G310100, a homolog of AtWRKY6, was identified
358 by (Pandey et al., 2011) as required for *Rpp2*-mediated resistance to ASR.
359 Glyma.13G310100 was induced by *Rpp1* silencing with a log2FC of 3.46.

360

361 DISCUSSION

362 The BPMV-based VIGS system has proven to be a powerful tool for analyzing
363 soybean genes involved in both the recognition event and downstream signaling that
364 occurs during the resistance response to *P. pachyrhizi*. Examples of this include the
365 identification of *Rpp4* (Meyer et al., 2009; Morales et al., 2013) and the functional

analysis of genes involved in *Rpp1*- and *Rpp2*-mediated responses (Pandey et al., 2011; Cooper et al., 2013). In this study, we employed VIGS to test candidate NBS-LRR genes identified through sequencing BACs spanning the *Rpp1* locus. The results presented here demonstrate that the soybean resistance gene responsible for the IR to *P. pachyrhizi* encodes an NBS-LRR protein with an N-terminal ULP1 domain belonging to the C48 peptidase family.

Based on our previous VIGS work with *Rpp4* (Meyer et al., 2009) and downstream *Rpp2* signaling components (Pandey et al., 2011), we expected that silencing *Rpp1* would yield a TAN phenotype with abundant uredinia, indicating complete susceptibility. However, our results suggest *Rpp1* silencing altered the IR, but not defense. There are three possible explanations for this unexpected phenotype. First, although silencing of the *Rpp1* candidate genes greatly reduced the levels of *Rpp1* candidate transcripts, VIGS is not absolute. Reduced levels of *Rpp1*, or components of the *Rpp1* signaling pathway, may condition a RB reaction rather than the IR. Expression levels of *CcRpp1*, a *P. pachyrhizi* resistance gene obtained from *Cajanus cajan* (pigeonpea), condition different degrees of resistance (Kawashima et al., 2016). Plants homozygous for *CcRpp1* were immune to *P. pachyrhizi* and displayed no visible symptoms. However, hemizygous plants displayed a RB type resistance. Expression analysis of the *CcRpp1* transgene revealed greater expression in homozygous plants, suggesting that expression levels influenced the efficacy of the transgene (Kawashima et al., 2016). Although *CcRpp1* is not orthologous to *GmRpp1*, it demonstrates that

expression levels of resistance genes can have a profound impact on the phenotype of resistance to soybean rust.

A second explanation for the observed RB phenotype in *Rpp1*-silenced plants is that an additional gene may condition resistance to *P. pachyrhizi* in PI 200492. McClean and Byth (1980) were the first to demonstrate that *Rpp1* segregated as a single locus in PI 200492. Similarly, Hyten et al. (2007) mapped *Rpp1* as a single locus in L85-2378 (*Rpp1*), an isoline developed from Williams 82 and PI 200492 (Bernard et al., 1991). Silencing of the candidate *Rpp1* genes in L85-2378 also compromised the IR (data not shown), suggesting no *R* genes outside the *Rpp1* locus contribute to rust resistance. However, the *Rpp1* locus does contain multiple tightly linked resistance genes. Chakraborty et al. (2009) mapped *Rpp1b* RB resistance in PI 594538A to a region spanning the *Rpp1* locus and recently Sahoo et al. (2017) mapped the *Rps12* (*Resistance to Phytophthora sojae* 12) gene in PI 399036 to a 368 kb region encompassing the *Rpp1* locus. Since silencing *R3*, *R4* and *R5* compromised the IR, but not defense to rust isolate LA04-1, a second *R* gene in the *Rpp1* locus would have to confer RB resistance to LA04-1 and not be silenced by our VIGS construct. This leaves *R6* and *R7* and additional *R* genes near the *Rpp1* locus not spanned by our BAC contig, as the most likely candidates for an additional rust resistance gene. VIGS constructs developed from identical regions in *R6* and *R7* failed to compromise resistance, even when co-silenced with *R3*, *R4* and *R5* (data not shown).

A third explanation is that *Rpp1* confers a completely novel phenotype. This hypothesis is supported by the findings of Cooper et al. (2013) who reported disruption

409 of *Rpp1*-mediated immunity following silencing of five different genes implicated in
 410 *Rpp1*-mediated defense. While silencing resulted in small lesions associated with a
 411 hypersensitive response and cell death, no sign of rust sporulation was detected,
 412 mirroring the *Rpp1*-silencing phenotype. Further, PI 200492 has been tested with over
 413 30 distinct *P. pachyrhizi* isolates representing geographical and temporal variation and
 414 only the IR or TAN phenotypes are observed (data not shown). Taken together these
 415 data suggest that no additional genes condition RB resistance to *P. pachyrhizi* in the
 416 *Rpp1* locus, but instead suggest silencing of *Rpp1* results in a novel phenotype. Future
 417 experiments involving *Rpp1* candidate transgene expression in a susceptible soybean
 418 background will hopefully shed light on these possibilities, demonstrating whether a
 419 single gene is sufficient to confer resistance and if so, whether the level of transgene
 420 expression influences the resistance phenotype. Furthermore, complementation with
 421 *Rpp1* candidate transgenes with inactive ULP1 domains should also provide evidence
 422 regarding the role of this domain in recognition and/or signaling.

423 The unusual structure of *Rpp1* raises several questions regarding how it
 424 functions in the resistance response. Specifically, does the ULP1 domain play a direct
 425 role in either recognition of the pathogen, downstream signaling following recognition,
 426 or potentially both events? Like other biotrophic plant-pathogenic fungi, *P. pachyrhizi*
 427 delivers immunity-suppressing effector proteins inside the plant cell via haustoria during
 428 the infection process to promote virulence. Little is known regarding the effectors
 429 deployed by *P. pachyrhizi*, but several transcriptomic studies have identified potential
 430 candidates that are likely transferred into the host (Link et al., 2014; Kunjeti et al., 2016;

431 de Carvalho et al., 2017). A small subset of these candidates have also been shown to
432 promote virulence (Kunjeti et al., 2016), suppress the HR (de Carvalho et al., 2017), or
433 suppress plant immunity (Qi et al., 2016; Qi et al., 2018). Although a cognate effector
434 protein that triggers *Rpp1*-mediated resistance has yet to be discovered, the Rpp1 ULP1
435 domain may serve as useful tool for the identification of at least one.

436 Plant NBS-LRR proteins have evolved to detect pathogen effectors inside the
437 host cell, either by direct or indirect recognition. In direct recognition, an effector
438 directly binds to the NBS-LRR protein, which triggers a response by the host cell. In
439 indirect recognition, the NBS-LRR protein monitors one or more host proteins acted
440 upon by the effector. If a host protein targeted by the effector plays a role in immunity
441 it is called a “guardee” (Dangl and Jones, 2001), and in cases where the host protein
442 mimics the authentic host protein targeted by the effector it is referred to as a “decoy”
443 (van der Hoorn and Kamoun, 2008). Recently, a new model that is an amalgam between
444 direct and indirect recognition has been put forth. This new model, known as the
445 “integrated decoy” model, attempts to explain why some NBS-LRR proteins contain
446 additional domains (Cesari et al., 2014). The authors of this model propose that some
447 host proteins targeted by effectors have been incorporated through evolution into NBS-
448 LRR proteins, allowing for direct recognition of effector proteins.

449 Although originally thought to be rather anomalous, the presence of additional
450 integrated domains within NBS-LRR proteins appears to be widespread throughout all
451 plant lineages (Cesari et al., 2013; Kroj et al., 2016; Sarris et al., 2016). In their search of
452 40 publically available plant predicted proteomes, Sarris et al. (2016) identified 61

distinct Pfam domains that occur in at least two different plant families. This group includes the ULP1 protease domain (called a Peptidase_C48 domain by Sarris et al. (2016)), which was found in NBS-LRR proteins in soybean, *Fragaria vesca*, and *Zea mays*. With the exception of soybean, presented in this study, these genes have not been demonstrated to play a role in defense. Nevertheless, it has been suggested that the mere presence of the same integrated domain in different plant species is an indication that these hosts may be attacked by pathogens that attempt to manipulate similar host proteins (Malik and Van der Hoorn, 2016). This hypothesis is consistent with the findings of Mukhtar et al. (2011) and Wessling et al. (2014) that demonstrate effectors from diverse pathogens may target common host proteins. We hypothesize that the ULP1 domain of Rpp1 functions as an integrated decoy and that one or more components of the sumoylation machinery may be targeted by *P. pachyrhizi* effectors.

The Rpp1 ULP1 domain most closely resembles proteases involved in sumoylation. Sumoylation is a rapid and reversible post-translational modification conserved in all eukaryotic organisms involving the conjugation of a SUMO peptide to a protein. The presence or absence of SUMO on a particular protein can alter its stability, interaction with other proteins, localization, or enzymatic activity, and there is an increasing body of evidence showing that sumoylation plays a key role in plant immunity (Verma et al., 2017). Among the first studies implicating sumoylation in plant-pathogen interactions was the discovery that XopD, a *Xanthomonas campestris* pv. *vesicatoria* effector protein, which is a SUMO protease with plant-specific SUMO substrate specificity (Hotson et al., 2003). During infection XopD is delivered inside the plant cell

475 via the bacterial type III secretion system where it mimics plant SUMO isopeptidases.
476 This finding demonstrates that pathogens likely target SUMO-dependent processes in
477 their attempts to thwart host defenses.

478 Forward and reverse genetic approaches have also identified components of the
479 SUMO machinery that participate in plant defense responses (Lee et al., 2007; van den
480 Burg and Takken, 2010; Bailey et al., 2016; Castaño-Miquel et al., 2017; Gou et al., 2017).
481 The *siz1* E3 SUMO ligase mutant in *Arabidopsis*, exhibits constitutive systemic-acquired
482 resistance (SAR) with elevated levels of salicylic acid (SA) and increased pathogenesis-
483 related (PR) genes (Lee et al., 2007). Interestingly, downregulation of soybean *GmSIZ1a*
484 and *GmSIZ1b* using RNA interference (RNAi)-mediated gene silencing did not
485 significantly alter SA levels (Cai et al., 2017), leaving open the possibility that other E3
486 SUMO ligases may play a role in SAR in soybean. NPR1, a central regulator of SAR, was
487 recently shown to be sumoylated upon immune induction by salicylic acid. Sumoylation
488 of NPR1 altered its association with WRKY transcriptional repressors to TGA
489 transcriptional activators (Saleh et al., 2015). Sumoylation of specific transcription
490 factors, including WRKYs, may also serve to suppress the expression of early defense
491 genes which are then rapidly activated through mitogen-activated protein kinase
492 (MAPK) signaling cascades in response to MTI or ETI (van den Burg and Takken, 2010).
493 Together, SUMO- and MAPK-signaling may provide dynamic control over plant immune
494 responses (van den Burg and Takken, 2010). Significantly, silencing of *GmNPR1* and
495 *GmWRKYs 36, 40, and 45* were shown to compromise *Rpp2*-mediated resistance
496 (Pandey et al., 2011).

497 Since independently evolved effectors from diverse pathogens target common
 498 host proteins, or “hub proteins” (Mukhtar et al., 2011; Wessling et al., 2014), it is
 499 reasonable to assume that enzymes involved in sumoylation, including the SUMO
 500 proteases that remove SUMO peptides from other proteins, are ideal targets for
 501 effectors that compromise MTI or ETI. Although we are unaware of any SUMO proteases
 502 that are targeted by plant pathogens, several fungal and oomycete effectors have been
 503 shown to exhibit protease inhibition (Tian et al., 2007; van Esse et al., 2008; Pretsch et
 504 al., 2013). RTP1 from the rust fungi *Uromyces fabae* and *U. striatus* were the first fungal
 505 proteins shown to be specifically expressed in haustoria and transferred to the host
 506 cytoplasm during infection (Kemen et al., 2005). More recently the C-terminal domain of
 507 RTP1 was shown to have similarity to the C-terminal domain of cysteine protease
 508 inhibitors (Pretsch et al., 2013). RTP1 homologs have now been identified in several rust
 509 fungi, including two from *P. pachyrhizi* (Pretsch et al., 2013). The putative targets of
 510 PpRTP1 and PpRTP2 are not known, but this indicates that *P. pachyrhizi* deploys
 511 protease inhibitors as part of its effector repertoire.

512 An additional inference that can be drawn from the integrated decoy model is
 513 that incorporated proteins may eventually lose their biochemical activity while retaining
 514 effector-binding properties (Maqbool et al., 2015). Analysis of protein kinases fused to
 515 NBS-LRR proteins, the most common class of integrated domain, revealed that most
 516 kinases identified as integrated domains are potentially catalytically active (Sarris et al.,
 517 2016). This suggests that integrated domains originating from proteins with enzymatic
 518 activity may not act merely as decoys; rather, such domains may retain their original

function and serve as integrated sensors (Sarris et al., 2016). Our analysis of the genes at the *Rpp1* locus is consistent with this assertion as the predicted proteins encoded by these genes all retain the conserved protease catalytic residues and expression of the ULP1 domain of R4 and R5 were shown to complement the yeast ULP mutants.

Further studies are needed to determine if the enzymatic activity of Rpp1 is required for the IR. If the protease activity is required, and assuming that disruption of the activity does not alter its interaction with a putative effector, then this would be an indication that the ULP1 domain functions in signal transduction. In this scenario, the ULP1 domain would become activated following the recognition event which would lead to the desumoylation of one or more downstream proteins leading to the IR. If the protease activity turns out to be dispensable for resistance, then the ULP1 domain of Rpp1 may have a function independent of resistance. The fact that the Williams 82 susceptible soybean line retains the ULP1 domains at this locus would argue that there may be a selective advantage to retaining this activity.

In previous experiments, we used microarray analysis to compare gene expression in control and *Rpp4*-silenced plants (Meyer et al., 2009). Unlike *Rpp1*, *Rpp4* contained no predicted integrated domains. Of the 264 differentially expressed genes responding to *Rpp4* silencing, 7.2% (19 of 264) were induced. In contrast, 32.1% (1,211 of 3,777) of differentially expressed genes were induced in response to *Rpp1*-silencing, suggesting the ULP1 domain found in R1, R3, R4 and R5 could impact gene expression. The *Arabidopsis* genome contains eight predicted ULPS: *OTS1/ULP1D* (At1g60220), *OTS2/ULP1C* (At1g10570), *ELS1/ULP1A* (At3g06910), *ESD4* (At4g15880),

541 *ASP1/SPF1/ULP2B* (At1g09730), *ULP1B* (At4g00690), *SPF2/ULP2A* (At4g33620) and
 542 At3g48480 (Reeves et al., 2002; Conti et al., 2008; Novatchkova et al., 2012; Kong et al.,
 543 2017). OTS1, OTS2, ESD4, ELS1 and ASP1 regulate flowering time in *Arabidopsis* (Reeves
 544 et al., 2002; Conti et al., 2008; Novatchkova et al., 2012; Kong et al., 2017). OTS1 and
 545 OTS2 regulate responses to osmotic stress, light-induced signaling and SA-mediated
 546 signaling (Sadanandom et al., 2015; Castro et al., 2016). In addition, OTS1 is involved in
 547 transcriptional gene silencing and tolerance to high copper levels (Liu et al., 2017a; Zhan
 548 et al., 2018). ASP1, which is most closely related to the ULP1 domains present in R1, R3,
 549 R4, and R5, regulates flowering time, fertility and ABA signaling during seedling
 550 development (Kong et al., 2017; Liu et al., 2017b; Wang et al., 2018). Of the eight
 551 *Arabidopsis* ULPs identified, only an OTS1 and OTS2 double mutant (*ulp1c/ulp1d*) has
 552 been characterized using whole genome expression analyses (Castro et al., 2016). In
 553 *ulp1c/ulp1d*, 112 genes were differentially expressed relative to wild-type controls.
 554 Genes associated with abiotic and biotic stress responses were significantly
 555 overrepresented. Nineteen of these genes (corresponding to 41 soybean best BLAST
 556 orthologs) were also differentially expressed in response to *Rpp1*-silencing. However
 557 none of the 41 soybean orthologs were differentially expressed in response to *Rpp4*-
 558 silencing, again suggesting a role for the ULP1 domain in R1, R3, R4 and R5. Included
 559 within the 19 genes common to *ulp1c/ulp1d* and *Rpp1*-silencing were orthologs of
 560 *Arabidopsis* flowering time genes (Flowering locus T (AtFT, Glyma.08G363100 and
 561 Glyma.16G150700) and AtFD (Glyma.04G022100), AtPSK5 (Glyma.09G277600 and
 562 Glyma.18G21270) and AtWRKY28 (Glyma.02G285900, Glyma.05G127600 and

563 Glyma.14G028900). AtPKS5 phosphorylates and interacts with NPR1 to regulate
564 expression of defense-related WRKYs WRKY33 and WRKY62 (Xie et al., 2010). AtWRKY28
565 enhances tolerance to abiotic and biotic stress (Babitha et al., 2013; Chen et al., 2013).
566 This suggests the ULP1 domain can modulate defense responses and helps explain the
567 novel phenotype of *Rpp1*-silenced plants.

568 Using a bioinformatic approach, Sarris et al. (2016) identified R proteins with
569 integrated domains in 40 different plant species. Interestingly, integration of the same
570 domain could be found in distinct plant lineages, indicating independent insertion
571 events. The peptidase C48 domain, corresponding to the ULP1 domain, was reported in
572 soybean, wild strawberry (*Fragaria vesca*, mrna24089.1-v1.0) and maize (*Zea mays*,
573 GRMZM2G033519_P01). In strawberry, the peptidase C48 domain was located
574 downstream of the LRR domain. In both soybean and maize, the peptidase C48 domain
575 was located upstream of the NBS. In maize, the peptidase C48 domain of
576 GRMZM2G033519_P01 contains multiple introns and exons. In contrast, the peptidase
577 C48 domain corresponds to a single exon in soybean. This confirms independent
578 integration of this domain in these three species. Given that all other ULP genes in the
579 soybean genome contain at least 4 exons (Li et al., 2017) and the presence of Copia
580 element immediately preceeding the ULP1 domain in *R1*, it seems likely the ULP1
581 domain was inserted into the first intron of an ancestral gene at the *Rpp1* locus by
582 readthrough transcription from a retroelement. Hoen et al. (2006) demonstrated that a
583 family of ULP-like genes (97 genes and psuedogenes) had expanded in the genome of
584 *Arabidopsis* through the action of Mutator-like transposable elements. Further,

transduplicated genes were under selective pressure to maintain their function. ULP domains have also been identified in transposable elements from diverse species including grape (Benjak et al., 2008), melon and rice (van Leeuwen et al., 2007) and from the soybean oomycete pathogens *Phytophthora sojae* and *P. ramorum* (Kojima and Jurka, 2011).

The identification of Rpp1 as a novel ULP1-NBS-LRR advances our understanding of the molecular basis of resistance employed by soybean against *P. pachyrhizi* and has allowed us to formulate new hypotheses to further explore this important pathosystem. For example, although many putative *P. pachyrhizi* effectors have been identified, little is known concerning their targets within the cell. The Rpp1 ULP1 domain may represent a key target of *P. pachyrhizi* effectors, which would provide additional support for the sumoylation pathway as an essential component of immunity towards microbial pathogens. Further investigation into the role of the Rpp1 ULP1 domain is an important future goal that should provide additional insight into the soybean defense mechanisms and *P. pachyrhizi* virulence determinants.

MATERIALS AND METHODS

Plant materials

The resistant soybean accession PI 200492 (Rpp1; McLean and Byth, 1980) and the susceptible genotype Williams 82 (Hyten et al., 2007) were used in this study. Seeds were germinated in a growth chamber at 20°C with a 16 h photoperiod. Plants

were fertilized with Peters Professional® 20-20-20 General Purpose (Everris NA Inc. Dublin, OH) at 3 weeks following germination.

Sequencing of the *Rpp1* locus

A bacterial artificial chromosome (BAC) library of the resistant soybean accession PI 200492 (*Rpp1*) was constructed from high molecular weight genomic DNA isolated from 10 g of young leaf tissue by Bio S&T Inc. (Montreal, Canada). The genomic DNA was partially digested with *HindIII*, cloned into pIndigo BAC (*HindIII*) (Epicentre Inc., Madison, WI, USA), and transformed into *E. coli* strain DH10B (Invitrogen, Canada). Estimated genome coverage was 10X with an average insert size of 130 Kb. The library was screened by PCR according to the manufacturer's recommendations with primers designed to amplify DNA sequences located at the *Rpp1* locus (Supplemental Table S1). Four BACs (PI_200492_1C06, PI_200492_2G03, PI_200492_1A02 and PI_200492_2E10) were identified and used to construct shotgun libraries. Briefly, DNA was extracted from *E. coli* using the Qiagen Large-Construct Kit (Qiagen, Valencia, CA, USA). The recovered DNA was sheared and cloned using the TOPO Shotgun Subcloning Kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA). DNA sequencing of subclones from BACs PI_200492_1A02 and PI_200492_2E10 was performed by Eurofins MWG Operon LLC (Huntsville, AL, USA). DNA sequencing of subclones from BACs PI_200492_1C06 and PI_200492_2G03 was performed in house at the USDA-ARS-CICGRU. All sequencing was performed using an Applied Biosystems 3730 DNA Analyzer with a 96-capillary array. Sequences were trimmed and assembled using Sequencher version 5.4 with the default parameters with the exception of a minimum match

percentage of 100% (Gene Codes Corporation) and a 100 basepair overlap. In order to maximize read lengths and improve assembly accuracy, forward and reverse reads for the same subclone were preassembled prior to BAC assembly. Each BAC was assembled individually, prior to assembly of all BACs. The Williams 82 reference genome (version Williams82.a2.v1, Schmutz et al., 2010) was used to orient contigs when necessary to fill gaps. Gaps were filled using polymerase chain reaction (PCR) to amplify from the BAC in question. PCR products were cloned and sequenced at the USDA-ARS-CICGRU facility using Hi-Fi Platinum Taq DNA polymerase (Invitrogen, no. 10342-053), TA cloning kit (Invitrogen, no. K4560-01), and One Shot TOP10 Electrocompetent *E. coli* (Invitrogen, no. C404052).

Annotation of the *Rpp1* locus in PI 200492

The sequence from the *Rpp1* contig from PI 200492 (324,316 bp, GenBank Accession MH590229) was divided into 2000 bp intervals which were compared to the UniProt (Apweiler et al., 2004) and TAIR (The *Arabidopsis* Information Resource version 10, www.arabidopsis.org) protein databases using BLASTX (Altschul et al., 1997) to identify potential protein coding sequences. In addition, BLASTN (Altschul et al., 1997) was used to compare against predicted genes and transcripts from the Williams 82 reference genome (Schmutz et al., 2010). FGENESH (Solovyev et al., 2006) and the NetPlantGene2 Server (Hebsgaard et al., 1996) analyses were used to predict exon positions and splice sites within candidate genes of interest. InterProScan (Jones et al., 2014) was used to identify conserved domains within predicted protein sequences of

649 interest. BLASTN (Altschul et al., 1997) against the SoyTE database (Du et al., 2010) was
 650 used to search for transposable elements within the *Rpp1* candidate genes.

651 **Evolutionary analyses of the *Rpp1* locus across legumes**

652 To examine the evolution of the *Rpp1* locus across sequenced legume species,
 653 we took advantage of the Legume Information System Genomic Context Viewer
 654 (https://legumeinfo.org/lis_context_viewer/, (Dash et al., 2016)). We used the sequence
 655 from the *Rpp1* locus in PI_200492 to identify the corresponding region in the Williams
 656 82 reference genome (Gm18 from 56.10 to 56.30 MB, version Wm82.a2.v1, (Schmutz et
 657 al., 2010)), the homeologous region on Gm08 (22.28 to 23.15 MB) of soybean, and
 658 syntenic regions from *Phaseolus vulgaris* (Pv08: 1.65 to 1.83 MB, genome version
 659 2.0, (Schmutz et al., 2014)), *Arachis duranensis* (A04: 120.49 to 121.05 MB, genome
 660 version 1.0, (Bertioli et al., 2015)), *Arachis ipaensis* (B04: 130.67 to 131.29 MB, genome
 661 version 1.0, (Bertioli et al., 2015)), *Cajanus cajan* (LG07: 16.70 to 16.79 MB, genome
 662 version 1.0, (Varshney et al., 2012)), *Lupinus angustifolius* (NLL08: 1.99 to 2.16 MB,
 663 genome version 1.0, (Hane et al., 2016)), *Lotus japonicus* (Chr01: 35.40 to 35.68 MB,
 664 genome version 2.5, (Sato et al., 2008)) and *Medicago truncatula* (Chr07: 5.93 to 6.60 MB,
 665 genome version 4.0, (Tang et al., 2014)). To determine if R proteins with ULP1 domains
 666 were present in any of these species, we divided each of the predicted protein
 667 sequences of R1, R3, R4 and R5 into two sections, one containing the ULP1 domain and
 668 the other containing the remaining protein sequence. BLASTP (Altschul et al., 1997) was
 669 used to compare each of these sections to all predicted proteins from soybean, *P.*
 670 *vulgaris*, *A. duranensis*, *A. ipaensis*, *C. cajan*, *L. angustifolius*, *L. japonicus* and *M.*

671 *truncatula*. BLAST reports of the two regions were compared to identify any putative R
 672 proteins with a ULP1 domain in any species.

673 ***P. pachyrhizi* inoculations**

674 *P. pachyrhizi* inoculum was prepared as previously described (Kendrick et al.,
 675 2011) using isolate LA04-1 which elicits an IR with soybean PI 200492 (*Rpp1*), but is
 676 fully pathogenic on the susceptible Williams 82 (Ray et al., 2009). Spore
 677 concentrations were measured with a hemacytometer and adjusted to approximately
 678 5.0×10^4 spores ml⁻¹ in a solution of 0.01% Tween 20 in sterile distilled water. Inoculum
 679 was applied with an atomizer until leaves were saturated. Plants were incubated in a
 680 20°C dew chamber overnight then transferred to a 25°C greenhouse.

681 **VIGS of *Rpp1* candidate genes**

682 A DNA fragment representing an identical 234 bp region of *R3*, *R4*, and *R5* was
 683 synthesized from overlapping oligonucleotides KP863, KP864, KP865, and KP866
 684 (Supplemental Table S1). Briefly, 2 µM of each oligonucleotide was reacted in PCR
 685 buffer with 0.2 mM dNTPs and 1 unit Taq polymerase with an initial denaturation step
 686 at 94°C for 2 min followed by 5 PCR cycles (94°C for 1 min, 50°C for 30 s, and 72°C for 1
 687 min) followed by 2 min extension at 72°C. One µl of the reaction was used as template
 688 for PCR under the same conditions with 20 cycles using oligonucleotides KP867 and
 689 KP868 to generate *Bam*HI and *Kpn*I sites for directional cloning into pBPMV-R2.
 690 Orientation of the insert was confirmed by sequencing using a vector-specific forward
 691 primer 1548F (Zhang et al., 2009). To generate inoculum for VIGS experiments, BPMV
 692 RNA1 (pBPMV-IA-R1M) and either the *Rpp1* experimental (pBPMV-*Rpp1*) or a green

fluorescent protein (GFP) control (pBPMV-*GFP*) plasmids were co-inoculated by via particle bombardment on leaves of Williams 82 plants at 14 days after sowing, as previously described (Zhang et al., 2009; 2010; Whitham et al., 2016). Symptomatic BPMV-infected leaf tissue was collected at 3 to 5 weeks after bombardment, lyophilized, and stored at -20°C .

Two weeks after germination, primary leaves of PI 200492 (*Rpp1*) plants were dusted with Carborundum and rub-inoculated with lyophilized BPMV-infected leaf tissue that was ground to a powder and suspended in 50 mM potassium phosphate buffer, pH 7.0. Plants were transferred to the United States Department of Agriculture BSL-3 plant pathogen containment facility at Fort Detrick (Melching et al., 1983) approximately 5 weeks after BPMV inoculation. BPMV-infected plants were inoculated with *P. pachyrhizi* and assessed for soybean rust symptoms 2 weeks after inoculation. For each experiment, six plants inoculated with each BPMV construct were used. Three independent replicates of the experiment were performed.

Assessment of gene silencing and fungal growth in VIGS plants

In order to assess *Rpp1* candidate gene silencing, 1 cm diameter leaf disks were removed from the fourth trifoliolate of BPMV-infected plants 2 weeks after inoculation with *P. pachyrhizi*. Collected tissue was immediately frozen in liquid nitrogen and stored at -80°C . Leaf tissue was ground in liquid nitrogen and RNA was extracted using the Qiagen Plant RNeasy kit (Qiagen, Valencia, CA, U.S.A.) and treated with Turbo DNase (Ambion, Thermo Fisher Scientific, Waltham, MA, U.S.A.). First-strand cDNA synthesis was performed with the Transcriptor First Strand cDNA synthesis kit (Roche,

Indianapolis, IN, U.S.A.). qRT-PCR was performed on cDNA derived from experimental and control plants using *Rpp1* candidate-specific oligonucleotide primers KP920 and KP921 and probe KP922 (Supplemental Table S1) modified with 6-carboxy fluorescein at the 5' end and with Blackhole Quencher I at the 3' end (Integrated DNA Technologies, Coralville, IA, U.S.A.) using a 120 s denaturation step at 95°C s followed by 40 PCR cycles (95°C for 15 s and 68°C for 60 s). Fungal growth was assessed by measuring the constitutively expressed *P. pachyrhizi* α -tubulin gene by qRT-PCR using primers KP1045 and KP1046 (Supplemental Table S1) using a 90 s denaturation step followed by 40 reaction cycles (95°C for 15 s, 60°C for 30 s, and 72°C for 30 s). Soybean *Rpp1* and *P. pachyrhizi* α -tubulin transcript levels were normalized to the soybean *ubiquitin-3* gene (GenBank accession no. D28123) using a 120 s denaturation at 95°C followed by 40 reaction cycles (95°C for 15 s, 52°C for 30 s, and 72°C for 30 s). The *ubiquitin-3* gene is not differentially expressed during *P. pachyrhizi* infection (van de Mortel et al., 2007). Three experimental replicates were performed for both qRT-PCR experiments using three independent biological replicates.

Gene expression of *Rpp1* candidate genes

To measure the expression of *Rpp1* candidate genes R1 through R5, we used ClustalW (Larkin et al., 2007) to align predicted RNA transcripts from PI 200492 and Williams 82. We identified an ~300bp region within the NBS domain that was highly conserved across all *R* genes in the *Rpp1* locus from PI 200492 and Williams 82 (Supplemental Figure S1). Primers (KP1041 and KP1042, Supplemental Table S1) were designed from regions identical across all genes. Amplified products varied in size and

737 contained unique sequence differences that could be used to identify the
738 corresponding genes. To measure amplification efficiency for each of the genes, the
739 primers were tested on genomic DNA of PI 200492 and Williams 82. Reaction
740 conditions included a 90 s denaturation step at 94°C followed by 35 PCR cycles (94°C
741 for 30 s, 54°C for 30 s, and 72°C for 60 s) and a 72°C extension for 120 s. PCR
742 amplicons were cloned into pCR2.1-TOPO (Invitrogen, Thermo Fisher Scientific,
743 Waltham, MA), and transformed into TOP10 Electrocomp *E. coli* (Invitrogen, Thermo
744 Fisher Scientific, Waltham, MA). Cloned DNA was prepared and sequenced at the USDA-
745 ARS-CICGRU as described above. Sequences were imported into Sequencher version 5.4
746 (Gene Codes Corporation, Ann Arbor, MI, USA) and vector sequences and low quality
747 sequences were trimmed. Sequences were aligned by clone pair and consensus
748 sequences were compared to the *Rpp1* locus in PI 200492 and Williams 82 using BLASTN
749 (Altschul et al., 1997) to remove non-target genes. Gene-specific nucleotide differences
750 were used to assign subclones to specific genes (Supplemental Figure S1 and Table S2).
751 Paired sequences of 100+ genomic clones from each genotype were used to determine
752 an application efficiency correction factor for *Rpp1* *R1* – *R5* in each genotype. Individual
753 gene representation was determined by dividing the number of clones assigned to each
754 gene by the total number of assigned clones overall. An amplification efficiency
755 correction factor was then determined for each gene in each genotype by dividing
756 expected representation by observed representation. Since the primers matched all
757 genes perfectly, and there were five genes in each genotype, expected representation
758 was 0.2 for all genes.

759 To measure baseline *Rpp1 R1 – R5* gene expression in resistant and susceptible
 760 plants, the second trifoliate leaves were collected from non-inoculated PI 200492 and
 761 Williams 82. To measure *Rpp1* candidate gene expression in response to *P. pachyrhizi*,
 762 tissue was collected from mock-inoculated and infected plants 24 and 72 hours after
 763 inoculation. Collected tissue was immediately frozen in liquid nitrogen and stored at
 764 -80°C . Leaf tissue was ground in liquid nitrogen and RNA was extracted using the
 765 Qiagen Plant RNeasy kit (Qiagen, Valencia, CA, U.S.A.) and treated with Turbo DNase
 766 (Ambion, Thermo Fisher Scientific, Waltham, MA, U.S.A.). First-strand cDNA synthesis
 767 was performed with the Transcriptor First Strand cDNA synthesis kit (Roche,
 768 Indianapolis, IN, U.S.A.). cDNAs were amplified with PCR primers KP1041 and KP1042,
 769 cloned, sequence and analyzed as described above. To correct expression for
 770 differences in amplification efficiency, the number of observed clones for a given cDNA
 771 was multiplied by the correction factor determined for each given gene above.

772 Experiments in yeast

773 Yeast strains TSA566 (MATa; *ura3 Δ 0*; *leu2 Δ 0*; *his3 Δ 1*; *met15 Δ 0*; *ulp1*-
 774 333:kanMX) and Y21424 (MATa/MAT α ; *ura3 Δ 0/ura3 Δ 0*; *leu2 Δ 0/leu2 Δ 0*; *his3 Δ 1/his3 Δ 1*;
 775 *met15 Δ 0/MET15*; *LYS2/lys2 Δ 0*; YIL031w/YIL031w::kanMX4) were obtained from the
 776 European *Saccharomyces cerevisiae* Archive for Functional Analysis (Euroscarf;
 777 www.euroscar.de). A haploid derivative of Y21424, AR01, was obtained via random
 778 spore analysis following standard protocols (Treco and Winston, 2008) and screened for
 779 lysine and methionine auxotrophic traits and the *Ulp2* deletion phenotype. Mutations in
 780 TSA566 (*ulp1*) and AR01 (*ulp2*) cause temperature sensitivity of growth. For

complementation studies TSA566 and AR01 were transformed with the ULP1 domains of *R3*, *R4* or *R5* driven by the galactose-inducible promoter P_{GAL1} or its weaker derivative P_{GALS} in the CEN6/URA3-based plasmids p416GAL1 and p416GALS, respectively (Mumberg et al., 1994). In addition, an inactive version of the *R4* ULP1 domain containing a mutation in the catalytic triad (C310S) was also expressed. Plasmids p416GAL1 and p416GALS were obtained from the American Type Culture Collection (ATCC; www.atcc.org). ORFs of the native *R3*, *R4* and *R5* ULP1 domains and mutant *R4* ULP1 domain, each modified to include an N-terminal myc tag were synthesized by GenScript (Piscataway, NJ) and excised from pUC57-Simple as *Xba*I/*Sal*I fragments and cloned into *Xba*I/*Sal*I digested p416GAL1 and p416GALS vectors. Transformed yeast cells were recovered on SD (–ura) with glucose as the sole carbon source at 30°C. To test for complementation, cells were grown overnight at 30°C in liquid SD (–ura) glucose media and streaked onto SD (–ura) glucose or SD (–ura) galactose/rafinose media followed by incubation at 30°C or 37°C for 3 days.

RNA-seq analyses of VIGS plants

The six RNA samples used to confirm VIGS silencing were also used for RNA-seq analyses. Prior to RNA-seq, samples were purified and concentrated using the Qiagen RNeasy MinuElute Cleanup Kit (74204, Qiagen, Germantown, MD). The Agilent 2100 BioanalyzerTM (Agilent, Santa Clara, CA) was used to confirm an RNA integrity number (RIN) greater than seven. RNA-seq library preparation and 150 bp single end sequencing was performed at the Iowa State University DNA facility, using the Illumina HiSeq 2500 platform (Illumina, San Diego, CA). Following sequencing, reads from individual libraries

803 were trimmed to remove adaptor sequences, sequencing artifacts and low quality
 804 sequences as described by Atwood et al. (Atwood et al., 2014). Tophat version 2.1.1
 805 (Trapnell et al., 2009) was used to align reads to the Williams 82 reference genome
 806 sequence (Wm82.a2.v1, (Schmutz et al., 2010)). Samtools (Li et al., 2009) was used to
 807 remove unreliably mapped reads. The resulting mapping files were imported in the
 808 statistical program R (R Core Team, 2014) using Rsamtools (Morgan and Pages, 2013).
 809 Gene correspondences were made using the Williams82.a2.v1 gene feature file,
 810 imported using rtracklayer (Lawrence et al., 2009). Mapped reads per gene per sample
 811 were counted using Genomic Alignments (Lawrence et al., 2013). The R graphics
 812 program ggplot2 (Wickham, 2009) was used to compare and visualize counts for sample
 813 replicates for technical reproducibility. One vector control sample was removed from
 814 further analysis.

815 Prior to data normalization and statistical analysis of differentially expressed
 816 genes, counts assigned *R3*, *R4* and *R5* were removed, as they were likely of viral origin.
 817 However, the high counts did confirm presence of the virus in silenced plants. All genes
 818 with counts per million <1 across all biological replicates were removed from further
 819 analysis. EdgeR (Robinson et al., 2010) was used for single factor, pairwise comparisons
 820 to calculate normalization factors and identify differentially expressed genes (DEGs)
 821 using a false discovery rate (FDR < .001). DEGs were annotated using the SoyBase
 822 Genome Annotation Report page (<https://www.soybase.org/genomeannotation/>) which
 823 provided best *A. thaliana* homologs and inferred gene ontology (GO) information (The
 824 *Arabidopsis* Information Resource [TAIR] version 10, www.arabidopsis.org). Significantly

(corrected P-value <0.05) overrepresented biological process GO terms were identified using the SoyBase GO Term Enrichment tool (https://www.soybase.org/goslimgraphic_v2/dashboard.php). Differentially expressed transcription factors were identified using The SoyDB transcription factor database (Wang et al., 2010).

ACKNOWLEDGEMENTS

We thank Andrea Luquette and Will Lane for their technical assistance. This research was financed by the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) Projects 8044-22000-046-00D and 5030-21220-005-00D, the Iowa Soybean Association, the North Central Soybean Research Program, the United Soybean Board and the National Science Foundation (award number 0820642).

AUTHOR CONTRIBUTIONS

K. Pedley, S. Whitham, and M. Graham designed the research; K. Pedley, A. Pandey, A. Ruck, L. Lincoln, and M. Graham planned and performed the experiments; all authors analyzed the data; K. Pedley and M. Graham wrote the manuscript; all authors read, edited, and approved the final manuscript.

844 **FIGURE LEGENDS**

845 **Figure 1.** Schematic diagram of the *Rpp1* locus. Four bacterial artificial chromosomes
846 (BACs; grey boxes) spanning the *Rpp1* locus on chromosome 18 were identified from a
847 library constructed from soybean accession PI 200492 (*Rpp1*) DNA. Positions of the
848 physical markers used to map *Rpp1* (Hyten et al., 2007), Sct_187 and Sat_046, are
849 illustrated at top along with targets of the primers used to screen the library. The blue
850 box indicates the location of the *Rpp1* locus. Primers were designed based on the
851 Williams 82 soybean genome sequence or PI_200492 BAC end sequences. The location
852 and orientation of the eight candidate *R* genes is depicted at bottom (green arrow
853 heads).

854
855 **Figure 2.** Structure of candidate *R* proteins in the *Rpp1* locus in PI 200492. Comparisons
856 to predicted genes and transcripts in the *Rpp1* locus from Williams 82 was used to
857 predict *R* gene structure in the *Rpp1* locus from PI 200492. The programs FGENESH
858 (Solovyev et al., 2006) and the NetPlantGene2 Server (Hebsgaard et al., 1996) were used
859 to confirm the positions of exons and splice sites within candidate *R* genes. InterProScan
860 (Jones et al., 2014) was used to identify conserved domains within *R* proteins. These
861 analyses divided the *R* proteins from the *Rpp1* locus into two distinct classes based on
862 motif differences and the presence or absence of introns in the corresponding genes.
863 Intron lengths (bp) are as follows: *R1* (4,485, 328 and 661), *R2* (635), *R3* (580, 329 and
864 712), *R4* (2,310, 332 and 674) and *R5* (1,153, 330 and 656).

865

Figure 3. Loss of *Rpp1*-mediated resistance in plants following VIGS of *Rpp1* candidate genes. Plants of soybean accession PI 200492 (*Rpp1*) were infected with *Bean pod mottle virus* (BPMV) carrying either a 234 bp fragment corresponding to an identical region in *R3*, *R4*, and *R5* or a portion of the green fluorescent protein gene (control). BPMV-infected plants were challenged with *P. pachyrhizi* isolate LA04-1 approximately 4 weeks after BPMV inoculation and symptoms were photographed 2 weeks later. *Rpp1*-silenced plants show typical soybean rust symptoms indicating a loss of immunity whereas control plants are immune. Top row images show the adaxial leaf surface and bottom row images show the abaxial leaf surface. Susceptible Williams 82 soybean plants (not infected with BPMV) were included to serve as a susceptible control. While *Rpp1*-silenced plants have RB lesions with few fungal spores, Williams 82 plants have tan lesions with fully sporulating uredinia.

Figure 4. Complementation of yeast mutants by *Rpp1* candidate gene ULP1 domains. The ULP1 domains of *R3*, *R4*, or *R5* were expressed from either a strong (Gal1) or weak (GalS) promoter in yeast mutants carrying either a temperature-sensitive (ts) allele of *ULP1* or lacking the *ULP2* gene. Both yeast mutants are normally impaired at 37°C. The *Rpp1* *R5* ULP1 domain was able to partially restore growth in the *ulp1 ts* mutant cells when expressed at high levels. In *ulp2* mutant cells, the *R4* ULP1 domain was able to complement the mutation at both high and low expression levels. Yeast cells were grown for 3 days and photographed.

888 **Figure 5.** Evolution of the *Rpp1* locus across legumes reveals the novel R proteins unique
 889 to soybean. The Legume Information System Genomic Context Viewer
 890 (https://legumeinfo.org/lis_context_viewer/, Dash et al. (2016)) was used to examine
 891 the region corresponding to *R1* through *R5* in the *Rpp1* locus (Gm18 from 56.10 to 56.30
 892 MB, version Wm82.a2.v1) across a broad range of legumes. The location and orientation
 893 of predicted genes is depicted by arrow heads. The color of the arrow head indicates
 894 sequence homology, a full description is provided in the box below the figure. To ease
 895 visualization, genes conserved across two or more species are aligned vertically.

896
 897 **Figure 6.** Characterization of the ULP1 domains in the *Rpp1* candidate genes. BLASTN
 898 comparing the ULP1 domains from *R1*, *R3*, *R4* and *R5* to all predicted transcripts in the
 899 Williams 82 reference genome identified best matches to homeologous genes
 900 Glyma.13G256800 ($E < 0.0$) and Glyma.15G058100 ($E < 10E-17$). While the ULP1 domain is
 901 encoded by a single exon in the *Rpp1* candidate genes, it corresponds to six exons in
 902 Glyma.13G256800 and Glyma.15G058100. Lack of introns in the ULP domains of the
 903 *Rpp1* candidate genes suggests the ULP1 domain was inserted into an ancestral *R* gene
 904 by a retrotransposition event between exons 1 and 2. BLASTN analyses identified a
 905 Copia retroelement in intron 1 of *R1*, immediately preceeding the ULP1 domain.

906
 907

908 **Supplemental Tables**

909

910 **Supplemental Table S1.** Oligonucleotides used in this study.

911

912 **Supplemental Table S2.** Cloning of genomic (DNA) and expression (cDNA) PCR products
913 from candidate R genes in the *Rpp1* locus in PI 200492 (*Rpp1*) and Williams 82.

914

915 **Supplemental Table S3.** Genes induced by *Rpp1*-silencing relative to vector control
916 plants.

917

918 **Supplemental Table S4.** Genes repressed by *GmRpp1*-silencing relative to vector control
919 plants.

920

921 **Supplemental Table S5.** Gene Ontology (GO) terms significantly (Corrected $P < .05$)
922 overrepresented among differentially expressed genes induced or repressed by *Rpp1*
923 silencing.

924

925 **Supplemental Table S6.** Transcription factors differentially expressed in response to
926 *GmRpp1* silencing.

927

928 **Supplemental Figures**

929 **Supplemental Figure S1.** Portion of NBS used for PCR, cloning and sequencing R gene
 930 products from the *Rpp1* locus in PI 200492 and Williams 82. Alignment was generated
 931 using CLUSTAL (Larkin et al., 2007) and bases are colored to indicate nucleotide identity
 932 or conservation. The location of primers KP1041 and KP1042 are indicated by black
 933 arrows.

934

935 **Supplemental Figure S2.** Sequence alignment of the ULP1 domains of candidate R
 936 proteins from the *Rpp1* locus in PI 200492 and Williams 82 with ULP proteases from
 937 *Arabidopsis thaliana* and *Saccharomyces cerevisiae*. Sequences were aligned using
 938 CLUSTAL (Larkin et al., 2007) and shaded using the CLUSTALX color scheme. Identical
 939 amino acids in the catalytic triad (His-Asp-Cys) are indicated with an asterisk.

940

941 **Supplemental Figure S3.** Complementation of yeast *ulp2* mutant with the R4 ULP1
 942 domain. The R4 ULP1 domain and a mutant version (C310S) of R4 were expressed using
 943 either a strong (Gal1) or weak (GalS) promoter in yeast cells lacking the *ULP2* gene.
 944 Growth of *ulp2* mutants are impaired at 37°C. The wild-type R4 ULP1 domain was able
 945 to restore growth in the *ulp2* mutant cells when expressed at high levels or low levels.
 946 The R4 (C310S) domain was unable to complement the mutation. Yeast cells were
 947 grown for 3 days and photographed.

948

949

Literature Cited

- 950
951
952 Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daniels, M.J., and Parker, J.E. 1998.
953 Different requirements for *EDS1* and *NDR1* by disease resistance genes define at
954 least two *R* gene-mediated signaling pathways in Arabidopsis. Proc. Natl. Acad.
955 Sci. U.S.A. 95:10306-10311.
- 956 Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman,
957 D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database
958 search programs. Nucleic Acids Res. 25:3389-3402.
- 959 Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E.,
960 Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C.,
961 Redaschi, N., and Yeh, L.S. 2004. UniProt: the universal protein knowledgebase.
962 Nucleic Acids Res. 32:D115-D119.
- 963 Atwood, S.E., O'Rourke, J.A., Peiffer, G.A., Yin, T., Majumder, M., Zhang, C., Cianzio, S.R.,
964 Hill, J.H., Cook, D., Whitham, S.A., Shoemaker, R.C., and Graham, M.A. 2014.
965 Replication protein A subunit 3 and the iron efficiency response in soybean. Plant
966 Cell Environ. 37:213-234.
- 967 Babitha, K., Ramu, S., Pruthvi, V., Mahesh, P., Nataraja, K.N., and Udayakumar, M. 2013.
968 Co-expression of AtbHLH17 and AtWRKY28 confers resistance to abiotic stress in
969 Arabidopsis. Transgenic Res. 22:327-341.
- 970 Bailey, M., Srivastava, A., Conti, L., Nelis, S., Zhang, C., Florance, H., Love, A., Milner, J.,
971 Napier, R., Grant, M., and Sadanandom, A. 2016. Stability of small ubiquitin-like
972 modifier (SUMO) proteases OVERLY TOLERANT TO SALT1 and -2 modulates

- 973 salicylic acid signalling and SUMO1/2 conjugation in *Arabidopsis thaliana*. J. Exp.
 974 Bot. 67:353-363.
- 975 Bakshi, M., and Oelmüller, R. 2014. WRKY transcription factors: Jack of many trades in
 976 plants. Plant Signal. Behav. 9:e27700.
- 977 Benjak, A., Forneck, A., and Casacuberta, J.M. 2008. Genome-wide analysis of the “cut-
 978 and-paste” transposons of grapevine. PLoS One 3:e3107.
- 979 Bernard, R.L., Nelson, R.L., and Cremeens, C.R. 1991. USDA soybean genetics collection:
 980 isoline collection. Soybean Genetics Newsletter 18:27-57.
- 981 Bertoli, D.J., Cannon, S.B., Froenicke, L., Huang, G., Farmer, A.D., Cannon, E.K.S., Liu, X.,
 982 Gao, D., Clevenger, J., Dash, S., Ren, L., Moretzsohn, M.C., Shirasawa, K., Huang,
 983 W., Vidigal, B., Abernathy, B., Chu, Y., Niederhuth, C.E., Umale, P., Araújo, A.C.G.,
 984 Kozik, A., Kim, K.D., Burow, M.D., Varshney, R.K., Wang, X., Zhang, X., Barkley, N.,
 985 Guimarães, P.M., Isobe, S., Guo, B., Liao, B., Stalker, H.T., Schmitz, R.J., Scheffler,
 986 B.E., Leal-Bertoli, S.C.M., Xun, X., Jackson, S.A., Michelmore, R., and Ozias-Akins,
 987 P. 2015. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the
 988 diploid ancestors of cultivated peanut. Nat. Genet 47:438.
- 989 Boller, T., and Felix, G. 2009. A renaissance of elicitors: perception of microbe-associated
 990 molecular patterns and danger signals by pattern-recognition receptors. Annu.
 991 Rev. Plant Biol. 60:379-406.
- 992 Bonde, M.R., Nester, S.E., Austin, C.N., Stone, C.L., Frederick, R.D., Hartman, G.L., and
 993 Miles, M.R. 2006. Evaluation of virulence of *Phakopsora pachyrhizi* and *P.*
 994 *meibomia* isolates. Plant Dis. 90:708-716.

- 995 Bromfield, K.R. 1984. Soybean Rust. Monogr. No. 11. American Phytopathological
996 Society, St. Paul, MN.
- 997 Cai, B., Kong, X., Zhong, C., Sun, S., Zhou, X.F., Jin, Y.H., Wang, Y., Li, X., Zhu, Z., and Jin,
998 J.B. 2017. SUMO E3 ligases GmSIZ1a and GmSIZ1b regulate vegetative growth in
999 soybean. *J. Integr. Plant Biol.* 59:2-14.
- 1000 Castaño-Miquel, L., Mas, A., Teixeira, I., Seguí, J., Perearnau, A., Thampi, B.N., Schapire,
1001 A.L., Rodrigo, N., La Verde, G., Manrique, S., Coca, M., and Lois, L.M. 2017.
1002 SUMOylation inhibition mediated by disruption of SUMO E1-E2 interactions
1003 confers plant susceptibility to necrotrophic fungal pathogens. *Mol. Plant* 10:709-
1004 720.
- 1005 Castro, P.H., Couto, D., Freitas, S., Verde, N., Macho, A.P., Huguet, S., Botella, M.A., Ruiz-
1006 Albert, J., Tavares, R.M., and Bejarano, E.R. 2016. SUMO proteases ULP1c and
1007 ULP1d are required for development and osmotic stress responses in *Arabidopsis*
1008 *thaliana*. *Plant Mol. Biol.* 92:143-159.
- 1009 Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P.N. 2014. A novel
1010 conserved mechanism for plant NLR protein pairs: the “integrated decoy”
1011 hypothesis. *Front. Plant Sci.* 5:606.
- 1012 Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A., Rivas, S., Alaux, L.,
1013 Kanzaki, H., and Okuyama, Y. 2013. The rice resistance protein pair RGA4/RGA5
1014 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct
1015 binding. *Plant Cell* 25:1463-1481.

- 1016 Chakraborty, N., Curley, J., Frederick, R.D., Hyten, D.L., Nelson, R.L., Hartman, G.L., and
 1017 Diers, B.W. 2009. Mapping and confirmation of a new allele from soybean PI
 1018 594538A conferring RB lesion-type resistance to soybean rust. *Crop Sci.* 49:783-
 1019 790.
- 1020 Chen, X., Liu, J., Lin, G., Wang, A., Wang, Z., and Lu, G. 2013. Overexpression of
 1021 *AtWRKY28* and *AtWRKY75* in *Arabidopsis* enhances resistance to oxalic acid and
 1022 *Sclerotinia sclerotiorum*. *Plant Cell Rep.* 32:1589-1599.
- 1023 Childs, S.P., King, Z.R., Walker, D.R., Harris, D.K., Pedley, K.F., Buck, J.W., Boerma, H.R.,
 1024 and Li, Z. 2017. Discovery of a seventh *Rpp* soybean rust resistance locus in
 1025 soybean accession PI 605823. *Theor. Appl. Genet.*:1-15.
- 1026 Conti, L., Price, G., O'Donnell, E., Schwessinger, B., Dominy, P., and Sadanandom, A.
 1027 2008. Small ubiquitin-like modifier proteases OVERLY TOLERANT TO SALT1 and -2
 1028 regulate salt stress responses in *Arabidopsis*. *Plant Cell* 20:2894-2908.
- 1029 Cooper, B., Campbell, K.B., McMahon, M.B., and Luster, D.G. 2013. Disruption of *Rpp1*-
 1030 mediated soybean rust immunity by virus-induced gene silencing. *Plant Signal.*
 1031 *Behav.* 8:e27543.
- 1032 Dangl, J.L., and Jones, J.D. 2001. Plant pathogens and integrated defence responses to
 1033 infection. *Nature* 411:826-833.
- 1034 Dash, S., Campbell, J.D., Cannon, E.K.S., Cleary, A.M., Huang, W., Kalberer, S.R.,
 1035 Karingula, V., Rice, A.G., Singh, J., Umale, P.E., Weeks, N.T., Wilkey, A.P., Farmer,
 1036 A.D., and Cannon, S.B. 2016. Legume information system (LegumeInfo.org): a key

- 1037 component of a set of federated data resources for the legume family. *Nucleic*
- 1038 *Acids Res.* 44:D1181–D1188.
- 1039 de Carvalho, M.C., Costa Nascimento, L., Darben, L.M., Polizel-Podanosqui, A.M., Lopes-
- 1040 Caitar, V.S., Qi, M., Rocha, C.S., Carazzolle, M.F., Kuwahara, M.K., Pereira, G.A.,
- 1041 Abdelnoor, R.V., Whitham, S.A., and Marcelino-Guimaraes, F.C. 2017. Prediction
- 1042 of the in planta *Phakopsora pachyrhizi* secretome and potential effector families.
- 1043 *Mol. Plant Pathol.* 18:363-377.
- 1044 Dodds, P.N., and Rathjen, J.P. 2010. Plant immunity: towards an integrated view of
- 1045 plant-pathogen interactions. *Nat. Rev. Genet.* 11:539-548.
- 1046 Du, J., Grant, D., Tian, Z., Nelson, R.T., Zhu, L., Shoemaker, R.C., and Ma, J. 2010.
- 1047 SoyTEdb: a comprehensive database of transposable elements in the soybean
- 1048 genome. *BMC Genomics* 11:113.
- 1049 Elmore, J.M., Lin, Z.J., and Coaker, G. 2011. Plant NB-LRR signaling: upstreams and
- 1050 downstreams. *Curr. Opin. Plant Biol.* 14:365-371.
- 1051 Flor, H.H. 1946. Genetics of pathogenicity in *Melampsora lini*. *J. Agric. Res.* 73:337-357.
- 1052 Garcia, A., Calvo, É.S., de Souza Kiihl, R.A., Harada, A., Hiromoto, D.M., and Vieira, L.G.E.
- 1053 2008. Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance
- 1054 genes: discovery of a novel locus and alleles. *Theor. Appl. Genet.* 117:545-553.
- 1055 Gou, M., Huang, Q., Qian, W., Zhang, Z., Jia, Z., and Hua, J. 2017. Sumoylation E3 ligase
- 1056 SIZ1 modulates plant immunity partly through the immune receptor gene *SNC1*
- 1057 in Arabidopsis. *Mol. Plant-Microbe Interact.* 30:334-342.

- 1058 Hane, J.K., Ming, Y., Kamphuis, L.G., Nelson, M.N., Garg, G., Atkins, C.A., Bayer, P.E.,
 1059 Bravo, A., Bringans, S., Cannon, S., Edwards, D., Foley, R., Gao, L., Harrison, M.J.,
 1060 Huang, W., Hurgobin, B., Li, S., Liu, C., McGrath, A., Morahan, G., Murray, J.,
 1061 Weller, J., Jian, J., and Singh, K.B. 2016. A comprehensive draft genome sequence
 1062 for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant–
 1063 microbe interactions and legume evolution. *Plant Biotechnol. J.* 15:318-330.
- 1064 Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P., and Brunak, S.
 1065 1996. Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local
 1066 and global sequence information. *Nucleic Acids Res.* 24:3439-3452.
- 1067 Hoen, D.R., Park, K.C., Elrouby, N., Yu, Z., Mohabir, N., Cowan, R.K., and Bureau, T.E.
 1068 2006. Transposon-mediated expansion and diversification of a family of *ULP*-like
 1069 genes. *Mol. Biol. Evol.* 23:1254-1268.
- 1070 Hotson, A., Chosed, R., Shu, H., Orth, K., and Mudgett, M.B. 2003. *Xanthomonas* type III
 1071 effector XopD targets SUMO-conjugated proteins in planta. *Mol. Microbiol.*
 1072 50:377-389.
- 1073 Hyten, D.L., Smith, J.R., Frederick, R.D., Tucker, M.L., Song, Q., and Cregan, P.B. 2009.
 1074 Bulk segregant analysis using the GoldenGate assay to locate the locus that
 1075 confers resistance to soybean rust in soybean. *Crop Sci.* 49:265-271.
- 1076 Hyten, D.L., Hartman, G.L., Nelson, R.L., Frederick, R.D., Concibido, V.C., Narvel, J.M.,
 1077 and Cregan, P.B. 2007. Map location of the *Rpp1* locus that confers resistance to
 1078 soybean rust in soybean. *Crop Sci.* 47:837-838.

- 1079 Jones, J.D., Vance, R.E., and Dangl, J.L. 2016. Intracellular innate immune surveillance
1080 devices in plants and animals. *Science* 354:aaf6395.
- 1081 Jones, J.D.G., and Dangl, J.L. 2006. The plant immune system. *Nature* 444:323-329.
- 1082 Jones, P., Binns, D., Chang, H., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J.,
1083 Mitchell, A., and Nuka, G. 2014. InterProScan 5: genome-scale protein function
1084 classification. *Bioinformatics* 30:1236-1240.
- 1085 Kawashima, C.G., Guimaraes, G.A., Nogueira, S.R., MacLean, D., Cook, D.R., Steuernagel,
1086 B., Baek, J., Bouyioukos, C., Melo Bdo, V., Tristao, G., de Oliveira, J.C., Rauscher,
1087 G., Mittal, S., Panichelli, L., Bacot, K., Johnson, E., Iyer, G., Tabor, G., Wulff, B.B.,
1088 Ward, E., Raidan, G.J., Broglie, K.E., Wu, G., van Esse, H.P., Jones, J.D., and
1089 Brommonschenkel, S.H. 2016. A pigeonpea gene confers resistance to Asian
1090 soybean rust in soybean. *Nat. Biotechnol.* 34:661-665.
- 1091 Kemen, E., Kemen, A.C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., and Voegelé,
1092 R.T. 2005. Identification of a protein from rust fungi transferred from haustoria
1093 into infected plant cells. *Mol. Plant-Microbe Interact.* 18:1130-1139.
- 1094 Kendrick, M.D., Harris, D.K., Ha, B.-K., Hyten, D.L., Cregan, P.B., Frederick, R.D., Boerma,
1095 H.R., and Pedley, K.F. 2011. Identification of a second Asian soybean rust
1096 resistance gene in Hyuuga soybean. *Phytopathology* 101:535-543.
- 1097 Kojima, K.K., and Jurka, J. 2011. Crypton transposons: identification of new diverse
1098 families and ancient domestication events. *Mobile DNA* 2:12.

- 1099 Kong, X., Luo, X., Qu, G.P., Liu, P., and Jin, J.B. 2017. Arabidopsis SUMO protease ASP1
 1100 positively regulates flowering time partially through regulating FLC stability. J.
 1101 Integr. Plant Biol. 59:15-29.
- 1102 Kroj, T., Chanclud, E., Michel - Romiti, C., Grand, X., and Morel, J.B. 2016. Integration of
 1103 decoy domains derived from protein targets of pathogen effectors into plant
 1104 immune receptors is widespread. New Phytol. 210:618-626.
- 1105 Kunjeti, S.G., Iyer, G., Johnson, E., Li, E., Broglie, K.E., Rauscher, G., and Rairdan, G.J.
 1106 2016. Identification of *Phakopsora pachyrhizi* candidate effectors with virulence
 1107 activity in a distantly related pathosystem. Front. Plant Sci. 7:269.
- 1108 Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H.,
 1109 Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and
 1110 Higgins, D.G. 2007. ClustalW and ClustalX version 2. Bioinformatics 23: 2947-
 1111 2948.
- 1112 Lawrence, M., Gentleman, R., and Carey, V. 2009. rtracklayer: an R package for
 1113 interfacing with genome browsers. Bioinformatics 25:1841-1842.
- 1114 Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan,
 1115 M.T., and Carey, V.J. 2013. Software for computing and annotating genomic
 1116 ranges. PLoS Comput. Biol. 9:e1003118.
- 1117 Lee, J., Nam, J., Park, H.C., Na, G., Miura, K., Jin, J.B., Yoo, C.Y., Baek, D., Kim, D.H., Jeong,
 1118 J.C., Kim, D., Lee, S.Y., Salt, D.E., Mengiste, T., Gong, Q., Ma, S., Bohnert, H.J.,
 1119 Kwak, S.S., Bressan, R.A., Hasegawa, P.M., and Yun, D.J. 2007. Salicylic acid-

- 1120 mediated innate immunity in *Arabidopsis* is regulated by SIZ1 SUMO E3 ligase.
- 1121 Plant J. 49:79-90.
- 1122 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis,
- 1123 G., Durbin, R., and Thousand Genome Project Data Processing Subgroup. 2009.
- 1124 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-
- 1125 2079.
- 1126 Li, S., Smith, J.R., Ray, J.D., and Frederick, R.D. 2012. Identification of a new soybean rust
- 1127 resistance gene in PI 567102B. Theor. Appl. Genet. 125:133-142.
- 1128 Li, S.J., and Hochstrasser, M. 1999. A new protease required for cell-cycle progression in
- 1129 yeast. Nature 398:246-251.
- 1130 Li, S.J., and Hochstrasser, M. 2000. The yeast *ULP2* (*SMT4*) gene encodes a novel
- 1131 protease specific for the ubiquitin-like Smt3 protein. Mol. Cell. Biol. 20:2367-
- 1132 2377.
- 1133 Li, Y., Wang, G., Xu, Z., Li, J., Sun, M., Guo, J., and Ji, W. 2017. Organization and
- 1134 regulation of soybean SUMOylation system under abiotic stress conditions.
- 1135 Front. Plant Sci. 8:1458.
- 1136 Link, T.I., Lang, P., Scheffler, B.E., Duke, M.V., Graham, M.A., Cooper, B., Tucker, M.L.,
- 1137 van de Mortel, M., Voegelé, R.T., Mendgen, K., Baum, T.J., and Whitham, S.A.
- 1138 2014. The haustorial transcriptomes of *Uromyces appendiculatus* and
- 1139 *Phakopsora pachyrhizi* and their candidate effector families. Mol. Plant Pathol.
- 1140 15:379-393.

- 1141 Liu, L., Yan, X., Kong, X., Zhao, Y., Gong, Z., Jin, J.B., and Guo, Y. 2017a. Transcriptional
1142 gene silencing maintained by OTS1 SUMO Protease requires a DNA-dependent
1143 polymerase V-dependent pathway. *Plant Physiol.* 173:655-667.
- 1144 Liu, L., Jiang, Y., Zhang, X., Wang, X., Wang, Y., Han, Y., Coupland, G., Jin, J.B., Searle, I.R.,
1145 and Fu, Y. 2017b. Two SUMO proteases SUMO PROTEASE RELATED TO FERTILITY
1146 1 and -2 are required for fertility. *Plant Physiol.*:DOI:10.1104/pp.1117.0021.
- 1147 Malik, S., and Van der Hoorn, R.A. 2016. Inspirational decoys: a new hunt for effector
1148 targets. *New Phytol.* 210:371-373.
- 1149 Maqbool, A., Saitoh, H., Franceschetti, M., Stevenson, C.E., Uemura, A., Kanzaki, H.,
1150 Kamoun, S., Terauchi, R., and Banfield, M.J. 2015. Structural basis of pathogen
1151 recognition by an integrated HMA domain in a plant NLR immune receptor. *eLife*
1152 4:e08709.
- 1153 Martin, G.B., Bogdanove, A.J., and Sessa, G. 2003. Understanding the functions of plant
1154 disease resistance proteins. *Annu. Rev. Plant Biol.* 54:23-61.
- 1155 McLean, R.J., and Byth, D.E. 1980. Inheritance of resistance to rust (*Phakopsora*
1156 *pachyrhizi*) in soybeans. *Aust. J. Agric. Res.* 31:951-956.
- 1157 Melching, J.S., Bromfield, K.R., and Kingsolver, C.H. 1983. The plant pathogen
1158 containment facility at Frederick, Maryland. *Plant Dis.* 67:717-722.
- 1159 Meyer, J.D.F., Silva, D.C.G., Yang, C., Pedley, K., Zhang, C., van de Mortel, M., Hill, J.H.,
1160 Shoemaker, R.C., Abdelnoor, R.V., Whitham, S.A., and Graham, M.A. 2009.
1161 Identification and analyses of candidate genes for *Rpp4*-mediated resistance to
1162 Asian soybean rust in soybean. *Plant Physiol.* 150:295-307.

- 1163 Miles, M., Bonde, M., Nester, S., Berner, D., Frederick, R., and Hartman, G. 2011.
 1164 Characterizing resistance to *Phakopsora pachyrhizi* in soybean. Plant Dis. 95:577-
 1165 581.
- 1166 Miles, M.R., Frederick, R.D., and Hartman, G.L. 2003. Is the U.S. soybean crop at risk?
 1167 APSnet feature, June 2003. American Phytopathological Society, St. Paul, MN.
 1168 Online publication.
- 1169 Morales, A.M.A.P., O'Rourke, J., van de Mortel, M., Schneider, K.T., Bancroft, T.J.,
 1170 Borém, A., Nelson, R.T., Nettleton, D., Baum, T.J., Shoemaker, R.C., Frederick,
 1171 R.D., Abdelnoor, R.V., Pedley, K.F., Whitham, S.A., and Graham, M.A. 2013.
 1172 Transcriptome analyses and virus induced gene silencing identify genes in the
 1173 *Rpp4*-mediated Asian soybean rust resistance pathway. Funct. Plant Biol.
 1174 40:1029-1047.
- 1175 Morgan, M., and Pages, H. (2013). Rsamtools: Binary alignment (BAM), variant call
 1176 (BCF), or tabix file import (R package version 1.12.4).
- 1177 Mukhtar, M.S., Carvunis, A.R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., Tasan,
 1178 M., Galli, M., Hao, T., Nishimura, M.T., Pevzner, S.J., Donovan, S.E., Ghamsari, L.,
 1179 Santhanam, B., Romero, V., Poulin, M.M., Gebreab, F., Gutierrez, B.J., Tam, S.,
 1180 Monachello, D., Boxem, M., Harbort, C.J., McDonald, N., Gai, L., Chen, H., He, Y.,
 1181 European Union Effectoromics, C., Vandenhaute, J., Roth, F.P., Hill, D.E., Ecker,
 1182 J.R., Vidal, M., Beynon, J., Braun, P., and Dangl, J.L. 2011. Independently evolved
 1183 virulence effectors converge onto hubs in a plant immune system network.
 1184 Science 333:596-601.

- 1185 Mumberg, D., R., M., and Funk, M. 1994. Regulatable promoters of *Saccharomyces*
 1186 *cerevisiae*: comparison of transcriptional activity and their use for heterologous
 1187 expression. Nucleic Acids Res. 22:5767–5768.
- 1188 Novatchkova, M., Tomanov, K., Hofmann, K., Stuible, H.-P., and Bachmair, A. 2012.
 1189 Update on sumoylation: defining core components of the plant SUMO
 1190 conjugation system by phylogenetic comparison. New Phytol. 195:23-31.
- 1191 Pandey, A.K., Yang, C., Zhang, C., Graham, M.A., Horstman, H.D., Lee, Y., Zabolina, O.A.,
 1192 Hill, J.H., Pedley, K.F., and Whitham, S.A. 2011. Functional analysis of the Asian
 1193 soybean rust resistance pathway mediated by *Rpp2*. Mol. Plant-Microbe Interact.
 1194 24:194-206.
- 1195 Phukan, U.J., Jeena, G.S., and Shukla, R.K. 2016. WRKY transcription factors: molecular
 1196 regulation and stress responses in plants. Front. Plant Sci. 7:260.
- 1197 Pretsch, K., Kemen, A., Kemen, E., Geiger, M., Mendgen, K., and Voegelé, R. 2013. The
 1198 rust transferred proteins-a new family of effector proteins exhibiting protease
 1199 inhibitor function. Mol. Plant Pathol. 14:96-107.
- 1200 Qi, M., Link, T.I., Müller, M., Hirschburger, D., Pudake, R.N., Pedley, K.F., Braun, E.,
 1201 Voegelé, R.T., Baum, T.J., and Whitham, S.A. 2016. A small cysteine-rich protein
 1202 from the Asian soybean rust fungus, *Phakopsora pachyrhizi*, suppresses plant
 1203 immunity. PLoS Pathog. 12:e1005827.
- 1204 Qi, M., Grayczyk, J.P., Seitz, J.M., Lee, Y., Link, T.I., Choi, D., Pedley, K.F., Voegelé, R.T.,
 1205 Baum, T.J., and Whitham, S.A. 2018. Suppression or activation of immune

- 1206 responses by predicted secreted proteins of the soybean rust pathogen
- 1207 *Phakopsora pachyrhizi*. Mol. Plant-Microbe Interact. 31:163-174.
- 1208 R Core Team. (2014). R: A language and environment for statistical computing. R
- 1209 Foundation for Statistical Computing, Vienna, Austria.
- 1210 Ray, J.D., Morel, W., Smith, J.R., Frederick, R.D., and Miles, M.R. 2009. Genetics and
- 1211 mapping of adult plant rust resistance in soybean PI 587886 and PI 587880A.
- 1212 Theor. Appl. Genet. 119:271-280.
- 1213 Reeves, P.H., Murtas, G., Dash, S., and Coupland, G. 2002. *Early in short days 4*, a
- 1214 mutation in Arabidopsis that causes early flowering and reduces the mRNA
- 1215 abundance of the floral repressor FLC. Development 129:5349-5361.
- 1216 Robinson, M.D., McCarthy, D.J., and Smyth, G.K. 2010. edgeR: a Bioconductor package
- 1217 for differential expression analysis of digital gene expression data. Bioinformatics
- 1218 26:139-140.
- 1219 Sadanandom, A., Ádám, É., Orosa, B., Viczián, A., Klose, C., Zhang, C., Josse, E.-M.,
- 1220 Kozma-Bognár, L., and Nagy, F. 2015. SUMOylation of phytochrome-B negatively
- 1221 regulates light-induced signaling in *Arabidopsis thaliana*. Proc. Natl. Acad. of Sci.
- 1222 U.S.A 112:11108-11113.
- 1223 Sahoo, D.K., Abeysekara, N.S., Cianzio, S.R., Robertson, A.E., and Bhattacharyya, M.K.
- 1224 2017. A novel *Phytophthora sojae* resistance *Rps12* gene mapped to a genomic
- 1225 region that contains several *Rps* genes. PloS One 12:e0169950.
- 1226 Saleh, A., Withers, J., Mohan, R., Marques, J., Gu, Y., Yan, S., Zavaliev, R., Nomoto, M.,
- 1227 Tada, Y., and Dong, X. 2015. Posttranslational modifications of the master

- 1228 transcriptional regulator NPR1 enable dynamic but tight control of plant immune
 1229 responses. *Cell Host Microbe* 18:169-182.
- 1230 Sarris, P.F., Cevik, V., Dagdas, G., Jones, J.D., and Krasileva, K.V. 2016. Comparative
 1231 analysis of plant immune receptor architectures uncovers host proteins likely
 1232 targeted by pathogens. *BMC Biol.* 14:8.
- 1233 Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., Sasamoto, S.,
 1234 Watanabe, A., Ono, A., Kawashima, K., Fujishiro, T., Katoh, M., Kohara, M.,
 1235 Kishida, Y., Minami, C., Nakayama, S., Nakazaki, N., Shimizu, Y., Shinpo, S.,
 1236 Takahashi, C., Wada, T., Yamada, M., Ohmido, N., Hayashi, M., Fukui, K., Baba, T.,
 1237 Nakamichi, T., Mori, H., and Tabata, S. 2008. Genome structure of the legume,
 1238 *Lotus Japonicus*. *DNA Res.* 15:227-239.
- 1239 Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., Jenkins, J.,
 1240 Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S.,
 1241 Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M.A., Chovatia, M., Gepts, P.,
 1242 Goodstein, D.M., Gonzales, M., Hellsten, U., Hyten, D.L., Jia, G., Kelly, J.D.,
 1243 Kudrna, D., Lee, R., Richard, M.M.S., Miklas, P.N., Osorno, J.M., Rodrigues, J.,
 1244 Thareau, V., Urrea, C.A., Wang, M., Yu, Y., Zhang, M., Wing, R.A., Cregan, P.B.,
 1245 Rokhsar, D.S., and Jackson, S.A. 2014. A reference genome for common bean and
 1246 genome-wide analysis of dual domestications. *Nat. Genet* 46:707.
- 1247 Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song,
 1248 Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G., Yu, Y., Sakurai, T.,
 1249 Umezawa, T., Bhattacharyya, M., Sandhu, D., Valliyodan, B., Lindquist, E., Peto,

- 1250 M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B.,
 1251 Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.,
 1252 Shinozaki, K., Nguyen, H., Wing, R., Cregan, P., Specht, J., Grimwood, J., Rokhsar,
 1253 D., Stacey, G., Shoemaker, R., and Jackson, S. 2010. Genome sequence of the
 1254 palaeopolyploid soybean. *Nature* 463:178-183.
- 1255 Schneider, K.T., Van de Mortel, M., Bancroft, T.J., Braun, E., Nettleton, D., Nelson, R.T.,
 1256 Frederick, R.D., Baum, T.J., Graham, M.A., and Whitham, S.A. 2011. Biphaseic
 1257 gene expression changes elicited by *Phakopsora pachyrhizi* in soybean correlate
 1258 with fungal penetration and haustoria formation. *Plant Physiol.* 157:355-371.
- 1259 Shao, Z.Q., Xue, J.Y., Wu, P., Zhang, Y.M., Wu, Y., Hang, Y.Y., Wang, B., and Chen, J.Q.
 1260 2016. Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich
 1261 repeat genes reveal three anciently diverged classes with distinct evolutionary
 1262 patterns. *Plant Physiol.* 170:2095-2109.
- 1263 Silva, D.C., Yamanaka, N., Brogin, R.L., Arias, C.A., Nepomuceno, A.L., Di Mauro, A.O.,
 1264 Pereira, S.S., Nogueira, L.M., Passianotto, A.L., and Abdelnoor, R.V. 2008.
 1265 Molecular mapping of two loci that confer resistance to Asian rust in soybean.
 1266 *Theor. Appl. Genet.* 117:57-63.
- 1267 Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. 2006. Automatic annotation of
 1268 eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 7:S10.
- 1269 Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L.,
 1270 Childs, K.L., Yandell, M., Gundlach, H., Mayer, K.F., Schwartz, D.C., and Town,

- 1271 C.D. 2014. An improved genome release (version Mt4.0) for the model legume
1272 *Medicago truncatula*. BMC Genomics 15:312.
- 1273 Tian, M., Win, J., Song, J., van der Hoorn, R., van der Knaap, E., and Kamoun, S. 2007. A
1274 *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like
1275 apoplastic protease. Plant Physiol. 143:364-377.
- 1276 Trapnell, C., Pachter, L., and Salzberg, S.L. 2009. TopHat: discovering splice junctions
1277 with RNA-Seq. Bioinformatics 25:1105-1111.
- 1278 Treco, D.A., and Winston, F. 2008. Growth and manipulation of yeast. Curr. Protoc. Mol.
1279 Biol. 82:13.12.11-13.12.12.
- 1280 van de Mortel, M., Recknor, J.C., Graham, M.A., Nettleton, D., Dittman, J.D., Nelson,
1281 R.T., Godoy, C.V., Abdelnoor, R.V., Almeida, A.M.R., Baum, T.J., and Whitham,
1282 S.A. 2007. Distinct biphasic mRNA changes in response to Asian soybean rust
1283 infection. Mol. Plant-Microbe Interact. 20:887-899.
- 1284 van den Burg, H.A., and Takken, F.L. 2010. SUMO-, MAPK-, and resistance protein-
1285 signaling converge at transcription complexes that regulate plant innate
1286 immunity. Plant Signal. Behav. 5:1597-1601.
- 1287 van der Hoorn, R.A., and Kamoun, S. 2008. From guard to decoy: a new model for
1288 perception of plant pathogen effectors. Plant Cell 20:2009-2017.
- 1289 van Esse, H.P., Van't Klooster, J.W., Bolton, M.D., Yadeta, K.A., van Baarlen, P., Boeren,
1290 S., Vervoort, J., de Wit, P.J., and Thomma, B.P. 2008. The *Cladosporium fulvum*
1291 virulence protein Avr2 inhibits host proteases required for basal defense. Plant
1292 Cell 20:1948-1963.

- 1293 van Leeuwen, H., Monfort, A., and Puigdomenech, P. 2007. *Mutator*-like elements
 1294 identified in melon, Arabidopsis and rice contain ULP1 protease domains. Mol.
 1295 Genet. Genomics 277:357-364.
- 1296 Varshney, R.K., Chen, W., Li, Y., Bharti, A.K., Saxena, R.K., Schlueter, J.A., Donoghue,
 1297 M.T.A., Azam, S., Fan, G., Whaley, A.M., Farmer, A.D., Sheridan, J., Iwata, A.,
 1298 Tuteja, R., Penmetsa, R.V., Wu, W., Upadhyaya, H.D., Yang, S., Shah, T., Saxena,
 1299 K.B., Michael, T., McCombie, W.R., Yang, B., Zhang, G., Yang, H., Wang, J.,
 1300 Spillane, J., Cook, D.R., May, G.D., Xu, X., and Jackson, S.A. 2012. Draft genome
 1301 sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor
 1302 farmers. Nat. Biotechnol. 30:83.
- 1303 Verma, V., Crolley, F., and Sadanandom, A. 2017. Fifty Shades of SUMO: its role in
 1304 immunity and at the fulcrum of growth - defense balance. Mol. Plant Pathol.
 1305 19:1537-1544.
- 1306 Wang, Q., Qu, G.P., Kong, X., Yan, Y., Li, J., and Jin, J.B. 2018. Arabidopsis SUMO protease
 1307 ASP1 positively regulates ABA signaling during early seedling development. J.
 1308 Integr. Plant Biol. <https://doi.org/10.1111/jipb.12669>.
- 1309 Wang, Z., Libault, M., Joshi, T., Valliyodan, B., Nguyen, H.T., Xu, D., Stacey, G., and
 1310 Cheng, J. 2010. SoyDB: a knowledge database of soybean transcription factors.
 1311 BMC Plant Biol. 10:14.
- 1312 Wessling, R., Eppe, P., Altmann, S., He, Y., Yang, L., Henz, S.R., McDonald, N., Wiley, K.,
 1313 Bader, K.C., Glasser, C., Mukhtar, M.S., Haigis, S., Ghamsari, L., Stephens, A.E.,
 1314 Ecker, J.R., Vidal, M., Jones, J.D., Mayer, K.F., Ver Loren van Themaat, E., Weigel,

- 1315 D., Schulze-Lefert, P., Dangl, J.L., Panstruga, R., and Braun, P. 2014. Convergent
1316 targeting of a common host protein-network by pathogen effectors from three
1317 kingdoms of life. *Cell Host Microbe* 16:364-375.
- 1318 Whitham, S.A., Lincoln, L.M., Chowda-Reddy, R.V., Dittman, J.D., O'Rourke, J.A., and
1319 Graham, M.A. 2016. Virus induced gene silencing and transient gene expression
1320 in soybean (*Glycine max*) using Bean Pod Mottle Virus infectious clones. *Curr.*
1321 *Protoc. Plant Biol.* 1:263-283.
- 1322 Wickham, H. 2009. ggplot2: Elegant graphics for data analysis. in: *Use R!*, Springer-
1323 Verlag, New York.
- 1324 Xie, C., Zhou, X., Deng, X., and Guo, Y. 2010. PKS5, a SNF1-related kinase, interacts with
1325 and phosphorylates NPR1, and modulates expression of *WRKY38* and *WRKY62*. *J.*
1326 *Genet. Genomics* 37:359-369.
- 1327 Zhan, E., Zhou, H., Li, S., Liu, L., Tan, T., and Lin, H. 2018. OTS1 - dependent
1328 deSUMOylation increases tolerance to high copper levels in *Arabidopsis*. *J.*
1329 *Integr. Plant Biol.* 60:310-322.
- 1330 Zhang, C., Yang, C., Whitham, S.A., and Hill, J.H. 2009. Development and use of an
1331 efficient DNA-based viral gene silencing vector for soybean. *Mol. Plant-Microbe*
1332 *Interact.* 22:123-131.
- 1333 Zhang, C., Bradshaw, J.D., Whitham, S.A., and Hill, J.H. 2010. The development of an
1334 efficient multipurpose Bean pod mottle virus viral vector set for foreign gene
1335 expression and RNA silencing. *Plant Physiol.* 153:52-65.
- 1336

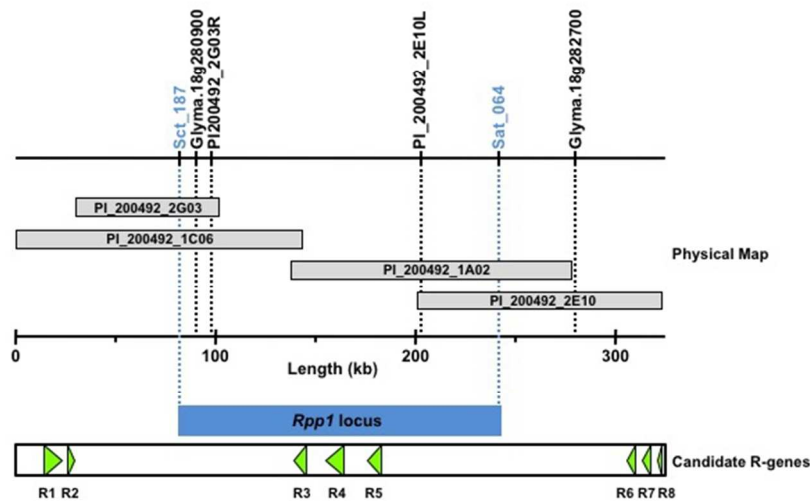


Figure 1. Pedley et al.

Figure 1. Schematic diagram of the *Rpp1* locus. Four bacterial artificial chromosomes (BACs; grey boxes) spanning the *Rpp1* locus on chromosome 18 were identified from a library constructed from soybean accession PI 200492 (*Rpp1*) DNA. Positions of the physical markers used to map *Rpp1* (Hyten et al., 2007), Sct_187 and Sat_046, are illustrated at top along with targets of the primers used to screen the library. The blue box indicates the location of the *Rpp1* locus. Primers were designed based on the Williams 82 soybean genome sequence or PI_200492 BAC end sequences. The location and orientation of the eight candidate *R* genes is depicted at bottom (green arrow heads).!! †

254x190mm (72 x 72 DPI)

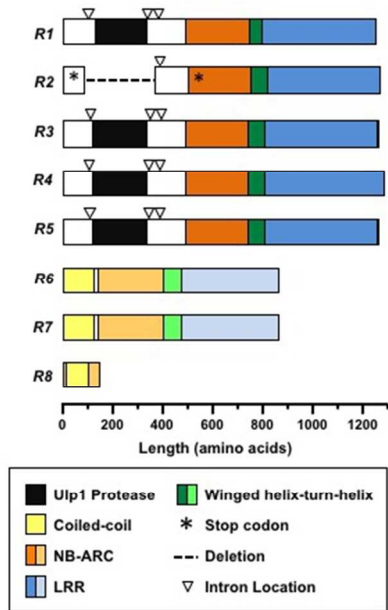


Figure 2. Pedley et al.

Figure 2. Structure of candidate R proteins in the *Rpp1* locus in PI 200492. Comparisons to predicted genes and transcripts in the *Rpp1* locus from Williams 82 was used to predict *R* gene structure in the *Rpp1* locus from PI 200492. The programs FGENESH (Solovyev et al., 2006) and the NetPlantGene2 Server (Hebsgaard et al., 1996) were used to confirm the positions of exons and splice sites within candidate *R* genes. InterProScan (Jones et al., 2014) was used to identify conserved domains within R proteins. These analyses divided the R proteins from the *Rpp1* locus into two distinct classes based on motif differences and the presence or absence of introns in the corresponding genes. Intron lengths (bp) are as follows: *R1* (4,485, 328 and 661), *R2* (635), *R3* (580, 329 and 712), *R4* (2,310, 332 and 674) and *R5* (1,153, 330 and 656).

254x190mm (72 x 72 DPI)

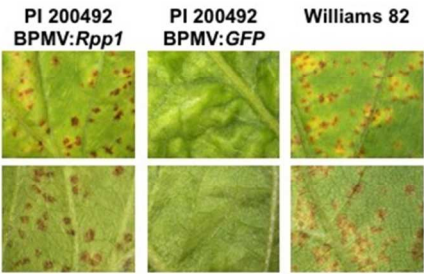


Figure 3. Pedley et al.

Figure 3. Loss of *Rpp1*-mediated resistance in plants following VIGS of *Rpp1* candidate genes. Plants of soybean accession PI 200492 (*Rpp1*) were infected with Bean pod mottle virus (BPMV) carrying either a 234 bp fragment corresponding to an identical region in *R3*, *R4*, and *R5* or a portion of the green fluorescent protein gene (control). BPMV-infected plants were challenged with *P. pachyrhizi* isolate LA04-1 approximately 4 weeks after BPMV inoculation and symptoms were photographed 2 weeks later. *Rpp1*-silenced plants show typical soybean rust symptoms indicating a loss of immunity whereas control plants are immune. Top row images show the adaxial leaf surface and bottom row images show the abaxial leaf surface. Susceptible Williams 82 soybean plants (not infected with BPMV) were included to serve as a susceptible control. While *Rpp1*-silenced plants have RB lesions with few fungal spores, Williams 82 plants have tan lesions with fully sporulating uredinia.

254x190mm (72 x 72 DPI)

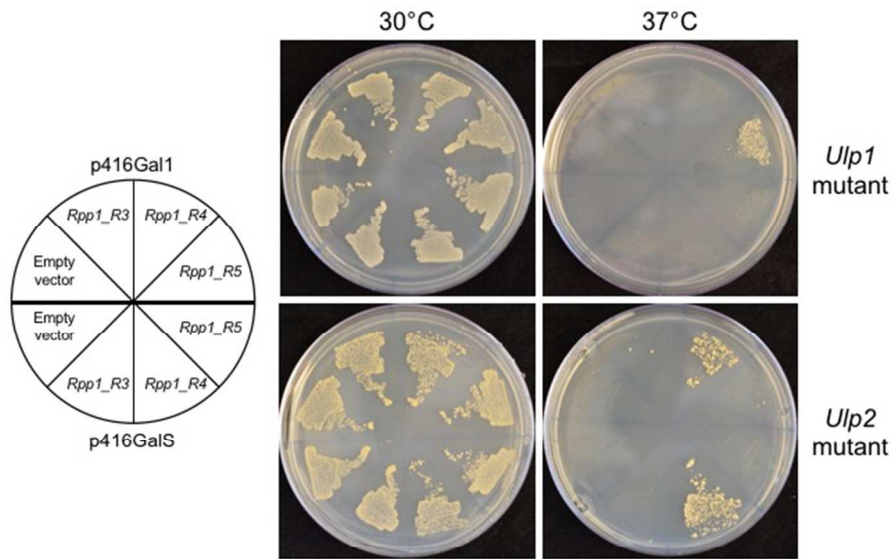


Figure 4. Pedley et al.

Figure 4. Complementation of yeast mutants by *Rpp1* candidate gene ULP1 domains. The ULP1 domains of *R3*, *R4*, or *R5* were expressed from either a strong (Gal1) or weak (GalS) promoter in yeast mutants carrying either a temperature-sensitive (*ts*) allele of *ULP1* or lacking the *ULP2* gene. Both yeast mutants are normally impaired at 37°C. The *Rpp1* *R5* ULP1 domain was able to partially restore growth in the *ulp1 ts* mutant cells when expressed at high levels. In *ulp2* mutant cells, the *R4* ULP1 domain was able to complement the mutation at both high and low expression levels. Yeast cells were grown for 3 days and photographed.

254x190mm (72 x 72 DPI)

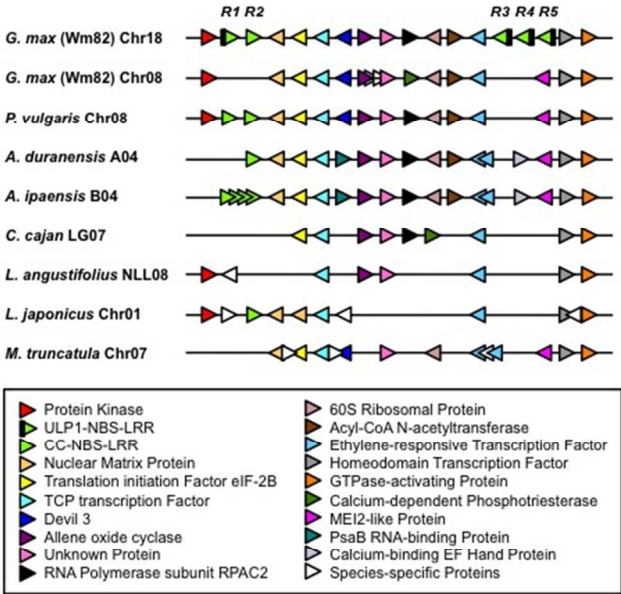


Figure 5. Pedley et al.

Figure 5. Evolution of the *Rpp1* locus across legumes reveals the novel R proteins unique to soybean. The Legume Information System Genomic Context Viewer (https://legumeinfo.org/lis_context_viewer/, Dash et al. (2016)) was used to examine the region corresponding to *R1* through *R5* in the *Rpp1* locus (Gm18 from 56.10 to 56.30 MB, version Wm82.a2.v1) across a broad range of legumes. The location and orientation of predicted genes is depicted by arrow heads. The color of the arrow head indicates sequence homology, a full description is provided in the box below the figure. To ease visualization, genes conserved across two or more species are aligned vertically.

254x190mm (72 x 72 DPI)

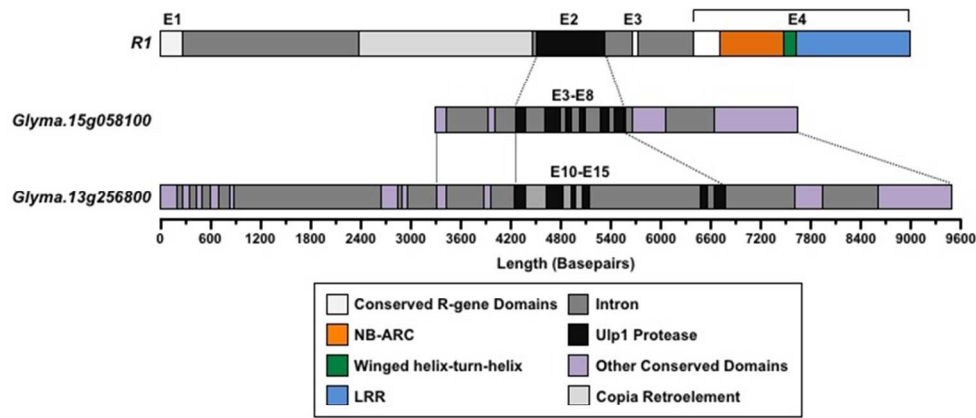
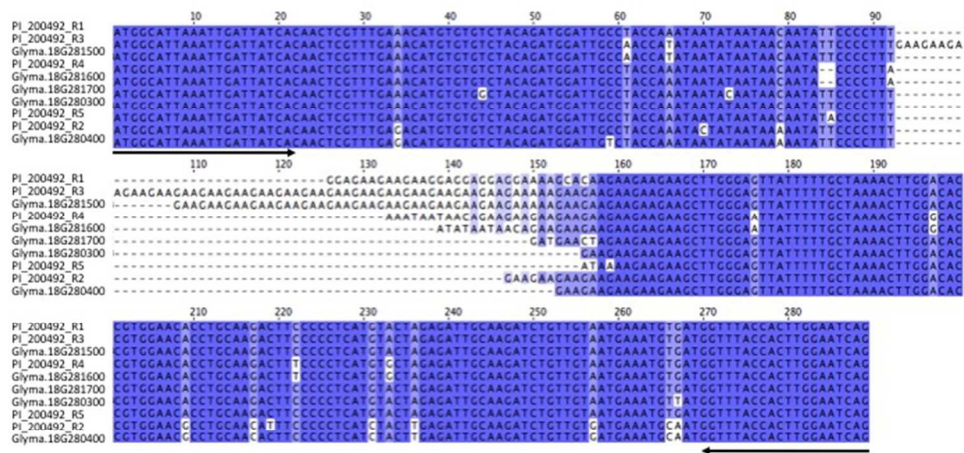


Figure 6. Pedley et al.

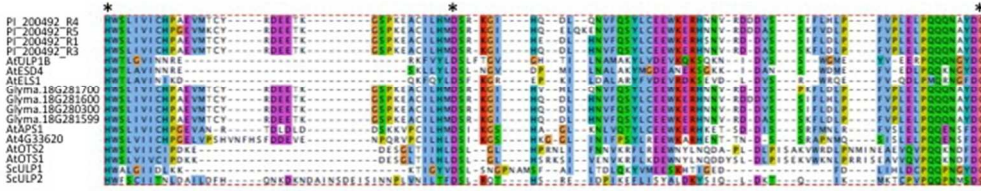
Figure 6. Characterization of the ULP1 domains in the *Rpp1* candidate genes. BLASTN comparing the ULP1 domains from *R1*, *R3*, *R4* and *R5* to all predicted transcripts in the Williams 82 reference genome identified best matches to homeologous genes *Glyma.13G256800* ($E < 0.0$) and *Glyma.15G058100* ($E < 10E-17$). While the ULP1 domain is encoded by a single exon in the *Rpp1* candidate genes, it corresponds to six exons in *Glyma.13G256800* and *Glyma.15G058100*. Lack of introns in the ULP domains of the *Rpp1* candidate genes suggests the ULP1 domain was inserted into an ancestral *R* gene by a retrotransposition event between exons 1 and 2. BLASTN analyses identified a Copia retroelement in intron 1 of *R1*, immediately preceding the ULP1 domain.

254x190mm (72 x 72 DPI)



Supplemental Figure S1. Portion of NBS used for PCR, cloning and sequencing R gene products from the *Rpp1* locus in PI 200492 and Williams 82. Alignment was generated using CLUSTAL (Larkin et al. 2007) and bases are colored to indicate nucleotide identity or conservation. The location of primers KP1041 and KP1042 are indicated by black arrows.

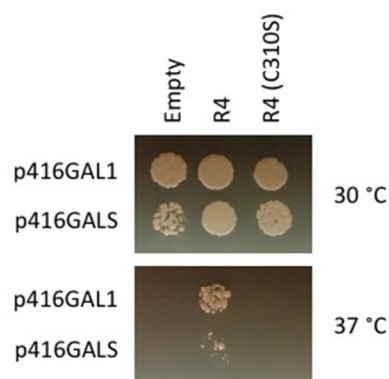
254x190mm (72 x 72 DPI)



Supplemental Figure S2. Pedley et al.

Supplemental Figure S2. Sequence alignment of the ULP1 domains of candidate R proteins from the *Rpp1* locus in PI 200492 and Williams 82 with ULP proteases from *Arabidopsis thaliana* and *Saccharomyces cerevisiae*. Sequences were aligned using CLUSTAL (Larkin et al. 2007) and shaded using the CLUSTALX color scheme. Identical amino acids in the catalytic triad (His-Asp-Cys) are indicated with an asterisk.

254x190mm (72 x 72 DPI)



Supplemental Figure S3. Pedley et al.

Supplemental Figure S3. Complementation of yeast *ulp2* mutant with the R4 ULP domain. The R4 ULP domain and a mutant version (C310S) of R4 were expressed from either a strong (Gal1) or weak (GalS) promoter in yeast cells lacking the *ULP2* gene. Growth of *ulp2* mutants are impaired at 37°C. The wild-type R4 ULP domain was able to restore growth in the *ulp2* mutant cells when expressed at high levels or low levels. The R4 (C310S) domain was unable to complement the mutation. Yeast cells were grown for 3 days and photographed.

254x190mm (72 x 72 DPI)